

8-1-2016

On Distributed and Acoustic Sensing for Situational Awareness

FANGRONG PENG

Syracuse University

Follow this and additional works at: <http://surface.syr.edu/etd>



Part of the [Engineering Commons](#)

Recommended Citation

PENG, FANGRONG, "On Distributed and Acoustic Sensing for Situational Awareness" (2016). *Dissertations - ALL*. 634.
<http://surface.syr.edu/etd/634>

This Dissertation is brought to you for free and open access by the SURFACE at SURFACE. It has been accepted for inclusion in Dissertations - ALL by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.

ABSTRACT

Recent advances in electronics enable the development of small-sized, low-cost, low-power, multi-functional sensor nodes that possess local processing capability as well as to work collaboratively through communications. They are able to sense, collect, and process data from the surrounding environment locally. Collaboration among the nodes are enabled due to their integrated communication capability. Such a system, generally referred to as sensor networks are widely used in various of areas, such as environmental monitoring, asset tracking, indoor navigation, etc.

This thesis consists of two separate applications of such mobile sensors. In this first part, we study decentralized inference problems with dependent observations in wireless sensor networks. Two separate problems are addressed in this part: one pertaining to collaborative spectrum sensing while the other on distributed parameter estimation with correlated additive Gaussian noise. In the second part, we employ a single acoustic sensor with co-located microphone and loudspeaker to reconstruct a 2-D convex polygonal room shape.

For spectrum sensing, we study the optimality of energy detection that has been widely used in the literature. This thesis studies the potential optimality (or sub-optimality) of the energy detector in spectrum sensing. With a single sensing node, we show that the energy detector is provably optimal for most cases and for the case when it is not theoretically optimal, its performance is nearly indistinguishable from the true optimal detector. For cooperative spectrum sensing where multiple nodes are employed, we use a recently proposed framework for distributed detection with dependent observations to establish the optimality of energy detector for several cooperative spectrum sensing systems and point out difficulties for the remaining cases.

The second problem in decentralized inference studied in this thesis is to investigate

the impact of noise correlation on decentralized estimation performance. For a tandem network with correlated additive Gaussian noises, we establish that threshold quantizer on local observations is optimal in the sense of maximizing Fisher information at the fusion center; this is true despite the fact that subsequent estimators may differ at the fusion center, depending on the statistical distribution of the parameter to be estimated. In addition, it is always beneficial to have the better sensor (i.e. the one with higher signal-to-noise ratio) serve as the fusion center in a tandem network for all correlation regimes. Finally, we identify different correlation regimes in terms of their impact on the estimation performance. These include the well known case where negatively correlated noises benefit estimation performance as it facilitates noise cancellation, as well as two distinct regimes with positively correlated noises compared with that of the independent case.

In the second part of this thesis, a practical problem of room shape reconstruction using first-order acoustic echoes is explored. Specifically, a single mobile node, with co-located loudspeaker, microphone and internal motion sensors, is deployed and times of arrival of the first-order echoes are measured and used to recover room shape. Two separate cases are studied: the first assumes no knowledge about the sensor trajectory, and the second one assumes partial knowledge on the sensor movement. For either case, the uniqueness of the mapping between the first-order echoes and the room geometry is discussed. Without any trajectory information, we show that first-order echoes are sufficient to recover 2-D room shapes for all convex polygons with the exception of parallelograms. Algorithmic procedure is developed to eliminate the higher-order echoes among the collected echoes in order to retrieve the room geometry. In the second case, the mapping is proved for any convex polygonal shapes when partial trajectory information from internal motion sensors is available. A practical algorithm for room reconstruction in the presence of noise and higher order echoes is proposed.

ON DISTRIBUTED AND ACOUSTIC SENSING FOR
SITUATIONAL AWARENESS

By

Fangrong Peng

B. S., Huazhong University of Science and Technology, Wuhan, China, 2008
M. S., Chalmers University of Technology, Gothenburg, Sweden, 2010

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in Electrical and Computer Engineering

Syracuse University
August 2016

Copyright © 2016 Fangrong Peng

All rights reserved

ACKNOWLEDGMENTS

First of all, I would like to express my gratitude to my advisor Dr. Biao Chen for his continued and unreserved supports during my doctoral studies, both intellectually and financially. It was a great pleasure to be a student of his. I sincerely thank him for his guidance on my research with his intellect, motivation, encouragement and patience, and the meetings with him makes me think more thoroughly and rigorously on my research topics. Without his help, I could not have finished this long and tough journey. I also want to thank Mrs. Chen for taking care of everyone in our lab, and she will always be deeply in my heart.

I would like to take this opportunity to thank my doctoral committee (alphabetically): Prof. Hao Chen, Prof. Pinyuen Chen, Prof. Yingbin Liang, Prof. Jian Tang, and Prof. Yanzhi Wang. Special thanks go to Dr. Pramod Washney and Dr. Makan Fardad, who are unfortunately not on my committee due to time conflict but provided valuable suggestions and encouragements to my defense. I would also like to thank my labmate Tiexing Wang for his collaboration in obtaining the results in Chapter 5 and Chapter 6. Thanks to Dr. Henk Wymeersch for leading me into the research world and providing me initial confidence. I am greatly indebted to my Communication Lab fellows, including Xiaohu, Yi, Wei, Minna, Ge, Fangfang, Kapil, Pengfei, Yu, Yang and Tiexing, with whom I shared my joy, and many even become my lifetime friends.

As a third-generation Ph.D. in my family, my parents thoroughly understand how tough the Ph.D. life is, and they offer me unconditional supports in every

aspect of my life. Special thanks to my father Dr. Fang Peng, for his inspiring advices when I was stuck in a mathematical proof. The thesis also goes in memory to my dearest grandmother, who passed away when I was absent from her, which left a deep hole in my heart permanently. I am also thankful to my beloved half Dr. Sijia Liu, with whom I go through my Ph.D. years also as friends and colleagues, and who makes me a better person.

I am grateful for my best friends in the United States, Dr. Yuting Zhang, Jiujiu Sun and Si Chen, with whom I began the relationships even back to high school, and from whom I still can get charged every time we hug each other. I am also obliged to my friends Yunyun Bi and Yue Duan for their genuine help during a hard time of my life.

Thanks to everyone who ever helped me and encouraged me, and who ever questioned me and challenged me.

*To my family:
past, present, and future.*

TABLE OF CONTENTS

Acknowledgments	v
List of Tables	xi
List of Figures	xii
1 Introduction	1
1.1 Decentralized Inference	1
1.1.1 Energy Detection in Spectrum Sensing	2
1.1.2 Data Dependency and Redundancy in a Tandem Network	3
1.2 Room Shape Reconstruction	4
2 Decentralized Inference with Dependent Observations	6
2.1 Bayesian Distributed Detection	8
2.2 Hierarchical Conditional Independence Model	10
2.2.1 CHCI model	11
3 Energy Detection in Cooperative Spectrum Sensing	13
3.1 Spectrum Sensing With A Single Node	14
3.1.1 Gaussian Signal over AWGN Channel	16
3.1.2 PSK Signal over Rayleigh Fading Channel	16
3.1.3 QAM Signal over Rayleigh Fading Channel	18
3.2 Cooperative Spectrum Sensing	21

3.2.1	Gaussian Signal over AWGN channel	24
3.2.2	PSK Signal over Fast Rayleigh Fading Channel	25
3.2.3	QAM Signal over Fast Rayleigh Fading Channel	26
3.3	Summary	27
4	Decentralized Estimation with Correlated Observations in a Tandem Network	28
4.1	Fisher Information	31
4.1.1	Classical Setting	31
4.1.2	Bayesian Framework	32
4.2	Centralized Estimation under Gaussian Noise	34
4.3	Decentralized Estimation with Bivariate Gaussian Noises	38
4.3.1	The Optimality of Single Threshold Quantizer	38
4.3.2	Communication Direction Problem	44
4.3.3	Data Dependency and Redundancy	46
4.4	Summary	52
5	Room Shape Recovery via a Single Mobile Acoustic Sensor without Trajectory	53
	Information	53
5.1	Problem Formulation	54
5.1.1	Image Source Model	54
5.1.2	Mapping between ISM and RIR	54
5.2	Room Shape Reconstruction	56
5.2.1	Identifiability by First-Order TOAs	56
5.2.2	Algorithm for Triangle Reconstruction with Only First-order TOAs	57
5.2.3	Polygon Reconstruction by First-Order TOAs	58
5.2.4	Algorithm of Polygon Reconstruction with Higher-order TOAs . . .	60
5.3	Summary	65

6	Room Shape Recovery via a Single Mobile Acoustic Sensor with Trajectory Information	66
6.1	System Model	67
6.1.1	Geometry	67
6.2	Recovery with Known Distances and Unknown Path Direction	69
6.3	Experimental result	73
6.3.1	Experiment Setup	73
6.3.2	Signal Type	74
6.3.3	Room Shape Reconstruction Experiment	74
6.4	Summary	75
7	Conclusion and Future Research Directions	77
7.1	Concluding Remarks	77
7.2	Future Research Directions	79
	Appendices	80
A	Distribution of PSK and QAM signals over Rayleigh fading channel	80
B	Proof of Lemma 4.3	82
C	Proof of Lemma 4.2	87
D	Centralized Estimation for a Multi-Node System in Gaussian Noises	90
E	Proof of Lemma 5.1	93
	References	96

LIST OF TABLES

LIST OF FIGURES

2.1	a canonical distributed detection system	7
3.1	ROC curves of the LRT and energy detector for 16QAM in Rayleigh fading. The SNR, defined in dB as $10 \log_{10} \frac{\bar{r}_m^2 \sigma^2}{\sigma_w^2}$, is at 5dB, and \bar{r}_m^2 is the average energy of QAM signals.	21
3.2	Deflection of two detectors within corresponding SNR ranges	22
4.1	Two fusion structures under a communication constraint where $\psi_i(\cdot)$ is assumed to be a one-bit quantizer in the present chapter	29
4.2	Comparison between $J(X; \theta)$ and $J(X; \theta Y)$	37
4.3	PCRLB vs. γ for different ρ	44
4.4	PCRLB comparison with $\rho = -0.5, 0, 0.5$	46
4.5	Comparison between $J(U(X); \theta)$ and $J(U(X); \theta Y)$	47
4.6	CRLB, its bounds and MSE of the Decentralized Estimation Problem	49
4.7	MSE vs. γ for different ρ	50
4.8	Estimation MSE comparison with $\rho = -0.5, 0, 0.5$	51
4.9	Estimation MSE with different SNR at the worse node	51
5.1	Image Source Model	55
5.2	original (blue) and recovered (dash purple/black) polygons: (a) an irregular quadrilateral, (b) a rotation, (c) a reflection, (d) a hexagon	64

6.1	A mobile device is employed to measure the geometry of a room. The mobile device collects echoes at O_1 , O_2 and O_3 successively. The distances between the consecutive measurement points are d_{12} and d_{23}	68
6.2	Correlator output at O_1 towards the first wall. Peaks with solid ellipses correspond to true walls. Peaks with dash ellipses correspond to either noise or higher-order echoes	74
6.3	Comparison between the ground truth (black underlined) and experiment result (red)	75
E.1	(a) one side is tilted, (b) two sides are tilted	94

CHAPTER 1

INTRODUCTION

This thesis consists of two parts. The first part focuses on decentralized inference with the emphasis on problems involving dependent observations across sensors. The second part deals with a practical research problem, in which we proposed computer-aided algorithms to reconstruct the shape of a room by using a single mobile acoustic sensor.

1.1 Decentralized Inference

Decentralized inference refers to the decision making process in a system where multiple distributed sensors are involved [1]. In such a system, each sensor observes a phenomenon of interest and transmits a compressed version of its observation to a fusion center (FC). Meanwhile, the FC makes final decision regarding the phenomenon using the information collected from the peripheral sensors, as well as its own observation. This general topic has drawn extensive attention over the past decades [2–6]. However, most of those works consider the scenario that the observations across sensors are statistically independent. When dependence occurs, the problem becomes much more challenging, as the local decision rule or quantizer design is usually coupled with each other. In this part of the thesis, we investigate two problems. The first one deals with decentralized detection; specifically, the

objective is to detect the presence of transmitted signals from the so-called primary user in the context of spectrum sensing using multiple nodes. The second one is on decentralized estimation where multiple sensors collectively estimate a parameter of interest. A common thread for both problems are that the observations across sensors are correlated with each other, both for the spectrum sensing and the decentralized estimation problems.

1.1.1 Energy Detection in Spectrum Sensing

We first study the problem of decentralized detection. Specifically, we analyze the energy detection in a cooperative spectrum sensing system in terms of its potential optimality. The spectrum sensing problem comes from the rising demand for accessing the wireless spectrums with the dramatically increasing number of wireless devices. Coincidentally, it has also long been recognized that there is gross under-utilization in licensed radio frequency bands [7]. In this case, dynamic spectrum access (DSA) has been proposed as one solution to resolve the spectrum crunch [8], which allows multiple secondary users to access spectrum space whenever the designated primary user is idle. A key enabling technology for DSA is spectrum sensing, i.e., a secondary user should only communicate when it believes that the primary user is indeed silent to avoid unintended interference to the primary user. However, many factors, including path-loss, shadowing, and channel fading, can degrade the sensing performance when only a single node is used for spectrum sensing. For example, the location of the secondary user may hinder its ability to hear the primary transmitter. To overcome this difficulty, cooperative spectrum sensing is introduced where spatially distributed nodes collaboratively detect the presence of the primary user's signal [9, 10].

While there have been numerous studies in spectrum sensing, it is almost without exception that an energy detector is used for both stand-alone spectrum sensing and cooperative spectrum sensing systems. It is somewhat surprising that there has not been any systematic study about the suitability or optimality of the energy detector for various spectrum sensing applications. Particularly, with random primary signals which are common

for all sensors, observations are correlated across sensing nodes given the hypothesis that the primary user is active. It is well known that for dependent observations, designing optimal local decision rules is generally an NP hard problem due to the coupling effect across sensors. This thesis fills this gap by analyzing energy detection in both stand-alone and cooperative spectrum sensing systems, thus providing a clear guidance on when such a detector may indeed be optimal for spectrum sensing.

1.1.2 Data Dependency and Redundancy in a Tandem Network

We then examine a problem within decentralized estimation system. The objective of distributed estimation is to estimate at the FC an underlying parameter which is indicative of the phenomenon of interest. Extensive studies have been reported for this topic during the decades (see [5, 6, 11–13] and references therein), and many fundamental results have been obtained. However, most results assume that the observations are conditionally independent (CI). We consider the challenging problem of decentralized estimation with dependent observations.

A decentralized estimation system usually requires quantization prior to the communications between the local sensors and the FC. This is due to various system limitations and resource constraints that collectively impose a finite capacity constraint. In the absence of CI assumption, the local quantizer design problem parallels that of local decision rule in decentralized detection system, which was known to be a NP hard problem. The work in this thesis provides some new insight on the optimum quantizer design for decentralized estimation with dependent observations. A better understanding of the optimum quantizer structure will in turn enable us to answer some of the interesting and important questions arising in decentralized inference. Two of them addressed in this dissertation include: optimal sensing architecture in terms of communication direction and the characterization of the correction regimes in terms of its impact on the estimation performance.

1.2 Room Shape Reconstruction

In the second part of this thesis, we study the indoor room shape recovery problem using a mobile acoustic sensor. Specifically, assuming a convex polygonal room and a mobile sensor with co-located loudspeaker and microphone, we study the problem of reconstructing the 2-D room geometry using the acoustic response of the room measured at multiple locations.

While reliable and accurate outdoor localization can be obtained by using primarily the global positioning system (GPS), significant obstacles exist to achieve similar localization performance in an indoor environment. GPS signals, due to its high frequencies, are often unavailable in an indoor environment due to severe attenuation, blockage, and other impairments. Instead, existing indoor localization often relies on pre-existing infrastructure such as WiFi signals, Bluetooth, ultra-wide band (UWB), LED light. For example, the technology introduced by WiFiSLAM (Simultaneous Localization and Mapping) requires the full coverage of WiFi signals for the indoor environment as well as pre-existing 2-D maps for learning and localization. Similarly, for UWB, LED light or Bluetooth based systems, certain anchor nodes need to be fixed in advance at some known locations in the indoor environment.

There are situations that the infrastructure is either unavailable or inaccessible. Even with applications where infrastructure may be pre-established, they may not be available in the event of natural disaster as power outage may render the infrastructure inaccessible. Therefore, being able to reconstruct the geometry of the surrounding environment and self-localize in the absence of pre-established infrastructure is critical for applications such as rescue missions by first responders.

Indoor room shape reconstruction using acoustic sensors has been an active field of research in recent years [14–19]. Most existing works, however, assume multiple loudspeaker and/or distributed microphones that simultaneously transmit and receive echoes within the room to be examined. In [14], a single loudspeaker and a microphone array

were used to measure multiple single wall impulse response in different angles. Then an $l-1$ regularized least squares method was applied to map the measured impulse responses to a 3D shoebox room. The number of walls was assumed to be known in this model. In [15], the time of arrival (TOA) was measured by a set of microphones, and an algorithm to eliminate the higher-order reflective signals (signals bouncing over more than one obstacles) and estimate the geometry was proposed, based on an interesting property of the Euclidean distance matrix. Without any prior information, especially the number of walls and the order of reflections, an TOA-based method was introduced in [16], where the TOAs were estimated by the generalized correlation method. A different model was proposed in [17], where multiple acoustic stimuli were used to generate echoes reflected under different angles to a single microphone. The reconstruction problem was addressed by finding common tangents of ellipses.

The work that is most closely related to our work is that of [18], where a single co-located loudspeaker and microphone is used for room shape recovery. Both first-order and second-order echoes are utilized in [18]. It was pointed out in [18] that first-order echoes alone are not sufficient for even the simplest room shape such as a triangle with a static sensor. This thesis introduces mobility of the acoustic sensor where multiple measurement points are used to collect acoustic echoes. Under minor assumption on the the measurement point (e.g., they are on a straight line) we establish that 2-D room shape reconstruction is indeed feasible using first order acoustic echoes when supplemental geometry information is available through internal motion sensor.

CHAPTER 2

DECENTRALIZED INFERENCE WITH DEPENDENT OBSERVATIONS

The first part of this thesis focus on decentralized inference with an emphasis on the problems containing dependent observations. In this chapter, we review a recently proposed technique tackling dependent observations [20] in the canonical distributed detection system as shown in Fig. 2.1.

In a canonical distributed detection network, each sensor observes a phenomenon distributed according to $p(x_1, \dots, x_K | \mathcal{H})$, and then makes a decision regarding the hypothesis \mathcal{H} , then the local decisions U_1, \dots, U_K are transmitted to the FC, where a final decision U_0 is made based on the received local decisions. Different from the centralized detection, the information accessible at the FC is the quantized version instead of the entirety of the original observations, usually due to the bandwidth limit. Extensive work has been done on this problem that leads to many fundamental results [2–4, 21–24].

In this system, there are two different design problems, the fusion rule and the local decision rule. The optimum fusion rule at the FC is known to be the likelihood ratio test (LRT) [22–24], while designing the local decision rule is much more complicated due to the coupling effect in the distributed setting. Most works assume that the local observations

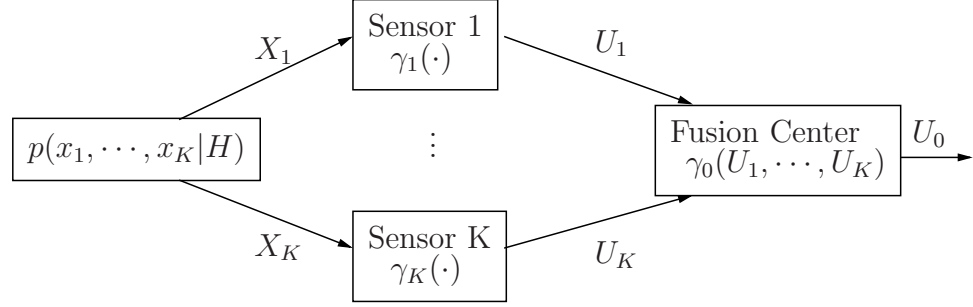


Fig. 2.1: a canonical distributed detection system

are conditionally independent of each other given each hypothesis, which implies that the joint distribution of the observations can be decomposed as

$$p(x_1, \dots, x_K | \mathcal{H}_i) = \prod_{k=1}^K p(x_k | \mathcal{H}_m), \quad m = 0, \dots, M - 1, \quad (2.1)$$

where M is the number of hypotheses in the detection problem.

Under this assumption, the problem can be significantly simplified. It is shown that the likelihood quantizer is optimal for the local decision rule [4, 25], while the coupling among the distributed sensors reduces to that on choosing the thresholds. In [2], a person-by-person optimization method is proposed to find the optimal quantization threshold for each local sensor.

In the case that the distributed sensors detect a random signal with independent noises or a deterministic signal embedded in correlated noises, the conditional independence (CI) assumption is violated, and the problem becomes generally intractable. It is shown in [26] that the local decision rule design is a NP complete problem when the observations are conditionally dependent, and the likelihood quantizer is generally not optimal, even for the relatively simple problem of two sensors observing a shift in mean correlated Gaussian random variables [27, 28].

A unifying model under the Bayesian inference framework is proposed in [20], and both conditionally independent and dependent observations can be considered as special cases of the new model. In such model, a hidden random variable \mathbf{W} is introduced such that the observations are conditionally independent on the new random variable, even if in the original model they are conditionally dependent given the underlying hypothesis. The new framework drastically simplifies the local decision rule design issue in a wide classes of distributed detection with dependent observations. The framework is described in the following section.

2.1 Bayesian Distributed Detection

Consider a parallel distributed hypothesis testing system with M hypotheses and K sensors illustrated in Fig. 2.1. The joint distribution of the local observations under each hypothesis $p(x_1, \dots, x_K|H)$, $H \in \{0, \dots, M-1\}$ is assumed to be known. Each sensor makes its local decision regarding the hypotheses based on its own observation X_k , and transmit the local decision $U_k = \gamma_k(X_k) \in \{0, \dots, M-1\}$ to the FC. Finally, the global decision $U_0 = \gamma_0(U_1, \dots, U_K) \in \{0, \dots, M-1\}$ is made by the FC. Let the prior distribution of the hypothesis H denoted as π_H , and the observations written in a vector form as $\mathbf{X} = \{X_1, \dots, X_K\}$. For simplicity, we denote $\{X_1, \dots, X_{k-1}, X_{k+1}, X_K\}$ as \mathbf{X}^k .

Generally, in the parallel distributed system, there is

$$p(\mathbf{u}|\mathbf{x}) = \prod_{k=1}^K p(u_k|x_k), \quad (2.2)$$

and the following Markov chain satisfies

$$H - \mathbf{X} - \mathbf{U} - U_0. \quad (2.3)$$

The goal in this parallel distributed detection system is to minimize the expected Bayesian

cost by choosing a set of decision rules $\{\gamma_0, \dots, \gamma_K\}$. Let $c_{u_0, h}$ be the Bayesian cost of deciding $U_0 = u_0$ when $H = h$ is true, and for measuring the probability of error, $c_{u_0, h}$ takes value from the set $\{0, 1\}$. Then the Bayesian cost of the whole system is

$$\begin{aligned} C &= \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(u_0|h) \\ &= \int_{\mathbf{X}} \sum_{\mathbf{u}} \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(u_0|\mathbf{u}) p(\mathbf{u}|\mathbf{x}) p(\mathbf{x}|h) d\mathbf{x}, \end{aligned} \quad (2.4)$$

where (2.4) follows from the Markov chain $H - \mathbf{X} - \mathbf{U} - U_0$. Then expand the Bayesian cost C with respect to the k -th sensor, we have

$$C = \int_{X_k} \sum_{u_k} p(u_k|x_k) f_k(u_k, x_k) dx_k, \quad (2.5)$$

where $f_k(u_k, x_k)$ is the Bayesian cost density function (BCDF) for the k th sensor making decision u_k while observing x_k , and it is defined as

$$\begin{aligned} f_k(u_k, x_k) &\triangleq \sum_{\mathbf{u}^k} \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(u_0|\mathbf{u}^k, u_k) \int_{\mathbf{X}^k} p(\mathbf{u}^k|\mathbf{x}^k) p(\mathbf{x}^k, x_k|h) d\mathbf{x}^k \\ &= \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(x_k|h) p(u_0|u_k, x_k, h), \end{aligned} \quad (2.6)$$

where (2.7) follows because

$$p(u_0|u_k, x_k, h) = \sum_{\mathbf{u}^k} p(u_0|\mathbf{u}^k, u_k) \int_{\mathbf{X}^k} p(\mathbf{u}^k|\mathbf{x}^k) p(\mathbf{x}^k|x_k, h) d\mathbf{x}^k.$$

From (2.5) one can see that to minimize the expected Bayesian cost, the k -th local sensor need to make a decision u_k such that $f_k(u_k, x_k)$ is minimized given fixed fusion rule and local decision rule at other sensors. On the other hand, since $f_k(u_k, x_k)$ is coupled with the fusion rule and other sensor's decision rule, finding the optimal fusion rule for the k -th sensor is generally intractable.

In the case that the observations are conditionally independent, the joint distribution can be decomposed according to (2.1), which leads the BCDF to

$$\begin{aligned}
f_k(u_k, x_k) &\triangleq \sum_{\mathbf{u}^k} \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(u_0 | \mathbf{u}^k, u_k) \int_{\mathbf{X}^k} p(\mathbf{u}^k | \mathbf{x}^k) p(\mathbf{x}^k | h) p(x_k | h) d\mathbf{x}^k \\
&= \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(x_k | h) p(u_0 | u_k, h), \\
&= \sum_{h=0}^{M-1} \alpha_k(u_k, h) p(x_k | h). \tag{2.7}
\end{aligned}$$

Therefore, the optimal decision rule $\gamma_k(X_k)$ reduces to an optimal M -ary Bayesian hypotheses test, i.e.

$$U_k = \gamma_k(X_k) = \arg \min_{u_k} \sum_{h=0}^{M-1} \alpha_k(u_k, h) p(x_k | h).$$

2.2 Hierarchical Conditional Independence Model

A hierarchical conditional independence (HCI) model [20] introduces a new random variable \mathbf{W} such that the following two conditions are satisfied,

1. the following Markov chain holds

$$H - \mathbf{W} - \mathbf{X} - \mathbf{U} - U_0.$$

2. X_1, \dots, X_K are conditionally independent given \mathbf{W} , i.e.

$$p(x_1, \dots, x_K | \mathbf{W}) = \prod_{k=1}^K p(x_k | \mathbf{W}).$$

Under this assumption, the joint distribution of local observations under each hypothesis can be rewritten as

$$p(x_1, \dots, x_K | H) = \int_{\mathbf{W}} p(\mathbf{w} | H) \prod_{k=1}^K p(x_k | \mathbf{w}). \quad (2.8)$$

Notice that the “hidden” variable can be scalar or vector. It has been shown that any distributed detection problem characterized by (2.3) can be represented by the HCI model [20]. The HCI model can be classified into three categories according to the support of \mathbf{W} , while in this chapter, we only focus on the continuous HCI (CHCI) model.

2.2.1 CHCI model

In this model, the conditional independence of local observations given the new random variable enables us to rewrite the BCDF:

$$\begin{aligned} f_k(u_k, x_k) &= \int_{\mathbf{W}} \sum_{u_0=0}^{M-1} \sum_{h=0}^{M-1} c_{u_0, h} \pi_h p(u_0 | u_k, w) p(w | h) p(x_k | w) dw \\ &= \int_{\mathbf{W}} \beta_{u_k, w} p(x_k | w) dw. \end{aligned} \quad (2.9)$$

It is shown in [20] that by imposing additional constraints on \mathbf{W} , a class of CHCI model can be determined where the optimal local decision rules are the threshold quantizers on local observations. The conclusion is repeated in the following proposition.

Proposition 2.1. [20, Proposition 1] *Consider a distributed detection system with binary hypothesis, binary sensor outputs, and the sensor observations are scalars. Suppose that the distributed detection problem can be described equivalently by the CHCI model where W is a scalar, and the following three conditions are satisfied:*

1. *The fusion center implements a monotone fusion rule, i.e.*

$$P(U_0 = 1 | U_k = 1, w) \geq P(U_0 = 1 | U_k = 0, w);$$

2. The ratio $\frac{p(w|\mathcal{H}_1)}{p(w|\mathcal{H}_0)}$ is a nondecreasing function of w ;
3. The ratio $\frac{p(x_k|w)}{p(x_k|w')}$ is also a nondecreasing function of w for any $x_k > x'_k$.

Then there exists a single threshold quantizer at the k th sensor

$$U_k = \begin{cases} 1 & x_k \geq \tau_k; \\ 0 & x_k < \tau_k, \end{cases}$$

that minimizes the error probability at the fusion center.

Proposition 2.1 serves as a new method to dealing with some distributed detection problems with conditionally dependent observations which seem formidable. The examples include the two shift in mean correlated Gaussian random variables [27, 28]. However, this proposition only works for the case that the sensor observations are scalars, and it will be seen in the next chapter that with vector observations, one needs to properly transform the observations before this proposition can be applied.

CHAPTER 3

ENERGY DETECTION IN COOPERATIVE SPECTRUM SENSING

The Energy detector was originally proposed for the detection of *an unknown deterministic signal* in additive white Gaussian noise (AWGN) [29]. The performance of energy detector of unknown deterministic signals over fading channels was analyzed in [30]. However, for spectrum sensing, it is often more convenient, and perhaps more accurate to treat the signals from primary user as random. Indeed, such assumption, i.e., digital signals are considered random is the very premise upon which the theory of optimum receiver in digital communications is developed. Additionally, the fading channels naturally randomize the transmitted signals even if one considers the signals themselves as deterministic. The assumption that the primary user's signal is of a random nature allows us to study the optimality of the energy detector.

In a spectrum sensing system, it keeps monitoring the spectrum occupancy to ensure the secondary users access the sparse portion without causing any undue interference to the primary user. In such a system, energy detection is heuristically used without validated optimality. The simplest model is to assume that the signal itself is Gaussian; if the channel itself is also assumed to be AWGN, an energy detector is provably optimal and

its performance can be analyzed in a straightforward manner [31]. The problem becomes much more difficult for cooperative spectrum sensing, even under the simple assumption of Gaussian signal over Gaussian channel. The presence of the common signal from the primary user introduces dependence among observations at different nodes and decentralized detection with dependent observations is a perennially difficult problem [28].

The chapter includes the detailed studies of the optimality or sub-optimality of the energy detector for various spectrum sensing systems. For single node spectrum sensing, we show that energy detection is provably optimal for detecting phase shift keying (PSK) signals in Rayleigh fading channels. For detecting quadrature amplitude modulation (QAM) in Rayleigh fading channels, energy detection is provably optimal for the low and high SNR regimes; for moderate SNR, while it is strictly sub-optimal, we show through numerical examples that its performance is practically identical to the optimum detector. For cooperative spectrum sensing, considerable challenges exist to obtain the optimality for the most general model. However, we are able to establish the optimality of energy detection for the following cases: Gaussian signals in Gaussian channels with a single observation; PSK signal in independent Rayleigh fading channels; QAM in independent Rayleigh fading channels with a single observation. Our results rely on the technique of tackling decentralized detection with dependent observations introduced in Chapter 2.

3.1 Spectrum Sensing With A Single Node

Consider the following hypothesis testing (HT) problem: for $n = 1, \dots, N$,

$$\begin{aligned}\mathcal{H}_0 : y(n) &= w(n) \\ \mathcal{H}_1 : y(n) &= x(n) + w(n),\end{aligned}\tag{3.1}$$

where n stands for the time index, and $w(n)$ is assumed throughout this chapter to be independent and identically distributed (i.i.d.) according to $\mathcal{CN}(0, \sigma_w^2)$, i.e., complex Gaussian

with zero mean and variance σ_w^2 . One can equivalently formulate the problem as testing whether or not $x(n) = 0$. The canonical form of an energy detector is the following test.

$$\sum_{n=1}^N |y(n)|^2 \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \lambda, \quad (3.2)$$

for some specified threshold λ . The energy detector was first proposed in [29] where $x(n)$ is assumed to be an unknown but deterministic signal. The detection performance can be derived by the distribution of the test statistic specified below.

$$\sum_{n=1}^N |y(n)|^2 \sim \begin{cases} \frac{\sigma_w^2}{2} \chi_{2N}^2 & \mathcal{H}_0; \\ \frac{\sigma_w^2}{2} \chi_{2N}^2(2snr) & \mathcal{H}_1, \end{cases}$$

where $2snr$ is the non-centrality parameter of the non-central chi-squared distribution, and $snr = \frac{\sum_{n=1}^N |x(n)|^2}{\sigma_w^2}$ is the signal to noise ratio (SNR) which depends on the total signal energy. Subsequently, the probability of detection (P_d) and probability of false alarm (P_f) can be obtained as

$$\begin{aligned} P_d &= P \left\{ \sum_{n=1}^N |y(n)|^2 > \lambda \mid \mathcal{H}_0 \right\} \\ &= Q_N \left(\sqrt{2snr}, \sqrt{\frac{\lambda}{\sigma_w^2/2}} \right), \\ P_f &= P \left\{ \sum_{n=1}^N |y(n)|^2 > \lambda \mid \mathcal{H}_1 \right\} \\ &= \frac{\Gamma \left(N, \frac{\lambda}{\sigma_w^2} \right)}{\Gamma(N)}, \end{aligned}$$

where $Q_N(\cdot, \cdot)$ is the generalized Marcum Q-function [32], representing the complement of the cumulative density function (CDF) of a non-central chi-squared distributed random variable, while $\Gamma(\cdot)$ is the gamma function, and $\Gamma(\cdot, \cdot)$ is the incomplete gamma function. The use of energy detection in detecting unknown deterministic signals in fading chan-

nel was further studied in [30] and the detection performance was analyzed under various fading scenarios.

In the following sections, we will instead consider $x(n)$ as a random signal, as is customary in the context of digital communications. We begin with the simple Gaussian signal model.

3.1.1 Gaussian Signal over AWGN Channel

By assuming that $x(n)$ is i.i.d. complex Gaussian [31] with zero mean and variance σ_s^2 , the HT problem in (3.1) becomes a classical random signal detection problem for which the energy detector can be easily shown to be the optimum LRT [33]. The distributions of the test statistic $\sum_{n=1}^N |y(n)|^2$ can be easily obtained to be [31]

$$\sum_{n=0}^N |y(n)|^2 \sim \begin{cases} \frac{\sigma_w^2}{2} \chi_{2N}^2 & \mathcal{H}_0; \\ \frac{\sigma_w^2 + \sigma_s^2}{2} \chi_{2N}^2 & \mathcal{H}_1. \end{cases}$$

Thus, instead of a test between a central chi-squared distribution against a non-central chi-squared distribution as in the case with deterministic signals, the problem becomes a test of two central chi-squared distributions with identical degrees of freedom but different scalings. The corresponding P_d and P_f are respectively

$$\begin{aligned} P_d &= \frac{\Gamma(N, \frac{\lambda}{\sigma_w^2 + \sigma_s^2})}{\Gamma(N)}, \\ P_f &= \frac{\Gamma(N, \frac{\lambda}{\sigma_w^2})}{\Gamma(N)}. \end{aligned} \tag{3.3}$$

3.1.2 PSK Signal over Rayleigh Fading Channel

Instead of AWGN channels, consider now that the primary signal is subject to a fading channel. For simplicity, we only consider Rayleigh flat fast fading in this thesis. As such,

the complex baseband signal in the HT problem (3.1) becomes

$$x(n) = s(n)h(n),$$

where $s(n)$ is a complex baseband signal while the fading coefficient $h(n)$ are zero-mean complex Gaussian variables with variance σ^2 whose amplitude follows a Rayleigh distribution. Furthermore, $h(n)$ are assumed to be independent of each other for different n , i.e., a fast fading channel. Let \mathcal{S} be the set of signal constellations for the primary user, the HT problem is to test the following two hypotheses: for $n = 1, \dots, N$,

$$\begin{aligned} \mathcal{H}_0 : y(n) &= w(n); \\ \mathcal{H}_1 : y(n) &= s(n)h(n) + w(n). \end{aligned} \quad (3.4)$$

In this subsection, we consider PSK modulation with M constellation points as the primary user's signal i.e. $s \in \mathcal{S}$, where

$$\mathcal{S} = \left\{ s_m : s_m = e^{j\theta_m}, \theta_m = m \cdot \frac{2\pi}{M} \text{ for } m = 1, \dots, M \right\} \quad (3.5)$$

For simplicity, s_m is assumed to be of unit energy as the signal energy can be easily absorbed into the fading channel statistics. Now let π_m be the prior probability for s_m , thus $\sum_{m=1}^M \pi_m = 1$. We show that $x(n)$ is also a complex Gaussian random variable. For ease of notation, we use $p(x)$ to denote $p_{x(n)}(x)$ in the following and similar for joint and conditional distributions. The distribution of the signal $x(n)$ is represented by

$$\begin{aligned} p(x) &= \sum_m p(x, s_m) \\ &= \sum_m \pi_m p(x|s_m) \\ &= \sum_m \pi_m p(h \cdot s_m) \end{aligned}$$

Given that h is circularly symmetric complex Gaussian distributed and that $s_m = e^{j\theta_m}$, the product $h \cdot s_m$ has exactly the same distribution as h for every m (rigorous proof see Appendix A). Thus the above mixture of densities gives rise to exactly the same distribution as h itself. It further implies that the signal $x(n) = h(n)s(n)$ observed at the sensing node is also a complex Gaussian signal. Since $x(n)$ is an independent Gaussian sequence across time where the independence comes from the independent fast fading assumption, $\sum_{n=1}^N |y(n)|^2$ is again the LRT for the HT problem. Hence an energy detector is optimal for the detection of PSK signals in independent Rayleigh fading channels.

3.1.3 QAM Signal over Rayleigh Fading Channel

Suppose instead that the set \mathcal{S} in (3.5) comes from a QAM constellation. Thus \mathcal{S} is defined as

$$\mathcal{S} = \{s_m : s_m = r_m e^{j\theta_m}, \text{ for } m = 1, \dots, M\}, \quad (3.6)$$

where r_m is not identical for all m and comes from a finite set related to the quantity M , while θ_m corresponds to each r_m . Again assign π_m as the prior probability to s_m . Similar to that for PSK, straightforward derivation shows that the density of $x(n)$ is a mixture of Gaussian densities but with *different* variances

$$p(y(n)|\mathcal{H}_1) = \sum_{m=1}^M \pi_m \cdot \frac{1}{\pi(r_m^2 \sigma^2 + \sigma_w^2)} e^{-\frac{|y(n)|^2}{r_m^2 \sigma^2 + \sigma_w^2}}$$

Apparently, it does not reduce further to a Gaussian variable as in the PSK case. Nevertheless, for a single sample, we can write out the likelihood ratio to be:

$$\begin{aligned} LR(y(n)) &= \frac{\sum_{m=1}^M \pi_m \cdot \frac{1}{\pi(r_m^2 \sigma^2 + \sigma_w^2)} \cdot e^{-\frac{|y(n)|^2}{r_m^2 \sigma^2 + \sigma_w^2}}}{\frac{1}{\pi \sigma_w^2} \cdot e^{-\frac{|y(n)|^2}{\sigma_w^2}}} \\ &= \sum_{m=1}^M \frac{\pi_m \cdot \sigma_w^2}{r_m^2 \sigma^2 + \sigma_w^2} \cdot e^{\frac{r_m^2 \sigma^2}{\sigma_w^2 (r_m^2 \sigma^2 + \sigma_w^2)} |y(n)|^2}, \end{aligned} \quad (3.7)$$

which is clearly a monotone increasing function of $|y(n)|^2$. Thus, if only a single sample (i.e., $N = 1$) is considered instead of a long sequence, the energy detector is again proved to be optimal.

For the cases of $N > 1$, under the independent fading assumption, the LR for N samples becomes:

$$LR(\mathbf{y}) = \prod_{n=1}^N \sum_{m=1}^M \frac{\pi_m \cdot \sigma_w^2}{r_m^2 \sigma^2 + \sigma_w^2} \cdot e^{\frac{r_m^2 \sigma^2}{\sigma_w^2 (r_m^2 \sigma^2 + \sigma_w^2)} \cdot |y(n)|^2}. \quad (3.8)$$

Apparently, the LR is not in general a monotone increasing function of $\sum_{n=1}^N |y(n)|^2$. Therefore, no general optimality of the energy detector can be claimed for such a case. However, we consider different regimes of SNR and show in the following that the energy detector is indeed asymptotically optimal in the high and low SNR regimes. The SNR in such fast fading channel is given by $\frac{r_m^2 \sigma^2}{\sigma_w^2}$.

1. High SNR regime.

Suppose

$$\frac{r_m^2 \sigma^2}{\sigma_w^2} \rightarrow \infty. \quad (3.9)$$

Thus (3.8) can be approximated as

$$\begin{aligned} LR(\mathbf{y}) &\approx \prod_{n=1}^N \sum_{m=1}^M \frac{\pi_m \sigma_w^2}{r_m^2 \sigma^2 + \sigma_w^2} e^{\frac{|y(n)|^2}{\sigma_w^2}} \\ &= \left(\sum_{m=1}^M \frac{\pi_m \sigma_w^2}{r_m^2 \sigma^2 + \sigma_w^2} \right)^N \prod_{n=1}^N e^{\frac{|y(n)|^2}{\sigma_w^2}} \\ &= \left(\sum_{m=1}^M \frac{\pi_m \sigma_w^2}{r_m^2 \sigma^2 + \sigma_w^2} \right)^N e^{\frac{1}{\sigma_w^2} \cdot \sum_{n=1}^N |y(n)|^2}, \end{aligned}$$

which is apparently a monotonically increasing function of the signal energy $\sum_{n=1}^N |y(n)|^2$.

Therefore, the energy detector is asymptotically optimal under the high SNR regime.

2. Low SNR regime.

Suppose instead

$$\frac{r_m^2 \sigma^2}{\sigma_w^2} \rightarrow 0. \quad (3.10)$$

With this assumption, (3.8) can be approximated as:

$$\begin{aligned} LR(\mathbf{y}) &\approx \prod_{n=1}^N \sum_{m=1}^M \pi_m e^{\frac{r_m^2 \sigma^2}{\sigma_w^4} \cdot |y(n)|^2} \\ &\approx \prod_{n=1}^N \sum_{m=1}^M \pi_m \left(1 + \frac{r_m^2 \sigma^2}{\sigma_w^4} \cdot |y(n)|^2 \right) \\ &= \prod_{n=1}^N \left(1 + \sum_{m=1}^M \frac{\pi_m r_m^2 \sigma^2}{\sigma_w^4} \cdot |y(n)|^2 \right). \end{aligned} \quad (3.11)$$

where we use the approximation $e^x \approx 1 + x$ for small x . Taking logarithm of (3.11) on both sides, the Log-Likelihood Ratio (LLR) is:

$$\begin{aligned} LLR(\mathbf{y}) &= \sum_{n=1}^N \log \left(1 + \sum_{m=1}^M \frac{\pi_m r_m^2 \sigma^2}{\sigma_w^4} \cdot |y(n)|^2 \right) \\ &\approx \sum_{n=1}^N \sum_{m=1}^M \frac{\pi_m r_m^2 \sigma^2}{\sigma_w^4} \cdot |y(n)|^2 \\ &= \sum_{m=1}^M \frac{\pi_m r_m^2 \sigma^2}{\sigma_w^4} \cdot \sum_{n=1}^N |y(n)|^2 \end{aligned}$$

where the approximation follows from $\log(1 + x) \approx x$ for small x . Again, for low SNR, the LRT reduces approximately to the energy detector.

3. Moderate SNR.

Besides of the two extreme cases discussed above, however, the energy detector is not provably optimal. Figs. 3.1, 3.2(a) and 3.2(b) are numerical examples comparing the LRT with that of the energy detector: the receiver operating characteristic (ROC) curve and the deflection measures [33] for the two test statistics. Fig. 3.1 is plotted at SNR=5 while Figs. 3.2(a) and 3.2(b) plot the deflection measures at different SNRs. In Fig. 3.2(a) the difference of the deflection measures between the two detectors

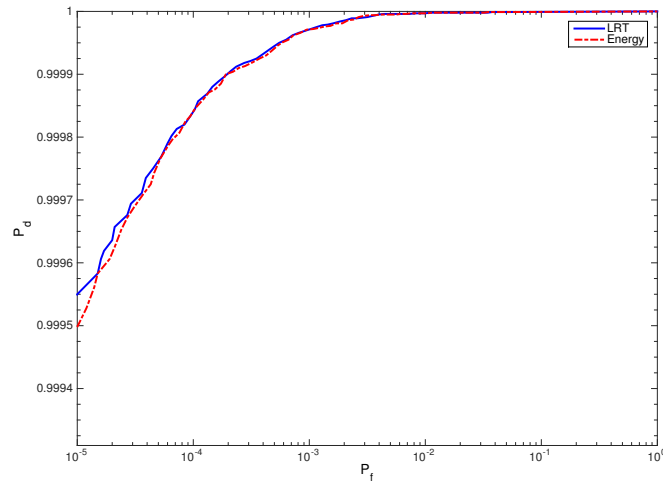


Fig. 3.1: ROC curves of the LRT and energy detector for 16QAM in Rayleigh fading. The SNR, defined in dB as $10 \log_{10} \frac{r_m^2 \sigma^2}{\sigma_w^2}$, is at 5dB, and r_m^2 is the average energy of QAM signals.

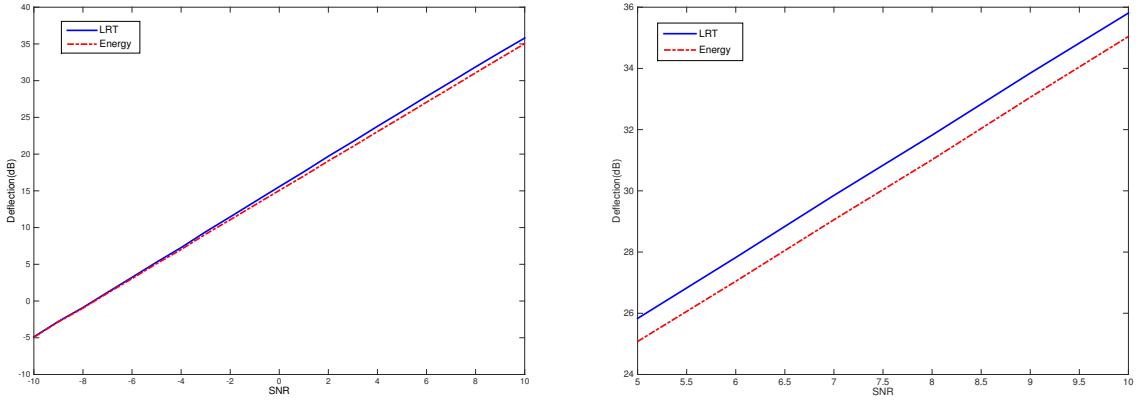
appears negligible; however, if we zoom in the curve in a particular interval, one can see noticeable difference in Fig. 3.2(b). LRT does outperform the energy detector *albeit* the difference is still small.

Thus for practical purposes, energy detector can be considered optimal for the QAM in Rayleigh fading case. From (3.8), the reason that the energy detector is nearly optimal can be attributed to the fact that the summation of different exponential terms is dominated by those points with large r_m . Thus the effect of unequal variances disappears if only the dominating terms are kept.

Summarizing, for single node spectrum sensing, energy detector is either provably optimal or close to optimal for all the cases investigated in this section.

3.2 Cooperative Spectrum Sensing

We now turn our attention to the case where spectrum sensing involves multiple nodes, i.e., cooperative spectrum sensing. Clearly, the problem of cooperative spectrum sensing



(a) Deflection of two detectors for SNR from -10 to 10 dB (b) Deflection of two detectors for SNR from 5 to 10 dB

Fig. 3.2: Deflection of two detectors within corresponding SNR ranges

becomes a classical decentralized detection problem. Suppose there are a total of K sensors (or nodes) in the system, cooperative spectrum sensing attempts to detect whether the primary user's signal is present or absent, i.e., to distinguish the following two hypotheses: for $k = 1, \dots, K$,

$$\begin{aligned} \mathcal{H}_0 : y_k(n) &= w_k(n), \\ \mathcal{H}_1 : y_k(n) &= x_k(n) + w_k(n), \end{aligned} \quad (3.12)$$

where $x_k(n)$ is the received signal at node k from the primary user and $w_k(n)$ is i.i.d. (in n) complex Gaussian noise for each k . Each sensor is assumed to make a binary decision, denoted by U_k , and report it to the fusion center where the final decision is made using U_1, \dots, U_K .

As with any decentralized detection problems, there are two different design issues: the fusion rule design and the decision rules at distributed sensors. The fusion rule design is a rather straightforward exercise given that it has input from all the sensors: if the fusion rule knows the local sensor decision rule, then optimal LRT can be easily implemented at the fusion center. The difficulty lies in the design of local decision rules as they are coupled with each other. Such coupling is especially acute when sensor observations are

conditionally dependent given one or both hypotheses. For such general cases, the optimal local sensor decision rule design is an NP-complete problem [26]. With conditionally independent observations, on the other hand, LRTs at local sensors are optimal for both the Bayesian and NP frameworks, and the coupling is only in terms of that the thresholds of the LRTs at different sensors are coupled [4, 34]. This significantly simplifies the problem as one can iteratively update, for each sensor, the LRT threshold; locally optimal thresholds can thus be easily attained.

For cooperative spectrum sensing, the corresponding decentralized detection problem has to deal with dependent observations: signals received at different sensors are generated by the same primary user, thus $x_k(n)$ are typically dependent of each other for different k under \mathcal{H}_1 . As such, even in cases where the energy detector coincides with the LRT at individual node, there is no guarantee that the energy detector is also optimal for the cooperative spectrum sensing system. In the following, we study several scenarios that parallel those cases studied in the previous section. Our approach in establishing the optimality of the energy detector for some of the cases largely relies on the results obtained in [35]. The new framework introduced in [35] assumes the existence of a hidden variable that induces conditional independence among otherwise dependent observations. Provided that the hidden variable satisfies certain conditions, then the optimal local sensor decision rules can be obtained in a straightforward manner. We repeat below Proposition 2.1 (also Proposition 1 in [35]) which is key to our approach. We denote $\mathbf{y}(n) \triangleq [y_1(n), y_2(n), \dots, y_K(n)]^T$ and tailor the proposition to our problem.

Proposition 3.1. *If there exists a scalar random variable W such that $\mathcal{H} - W - \mathbf{y}(n)$ forms a Markov chain, and it satisfies the following conditions:*

1. *The fusion center implements a monotone fusion rule;*
2. *The ratio $\frac{p(w|\mathcal{H}_1)}{p(w|\mathcal{H}_0)}$ is a nondecreasing function of w ;*
3. *The ratio $\frac{p(y_k(n)|w)}{p(y'_k(n)|w)}$ is also a nondecreasing function of w for any $y_k(n) > y'_k(n)$.*

Then there exists a single threshold quantizer at the k th sensor

$$U_k = \begin{cases} 1 & y_k(n) \geq \tau_k; \\ 0 & y_k(n) < \tau_k, \end{cases}$$

that minimizes the error probability at the fusion center.

We discuss in the following the Gaussian signal over the AWGN channel and the PSK as well as QAM under Rayleigh fading channel in the cooperative spectrum sensing system.

3.2.1 Gaussian Signal over AWGN channel

Given the model in (3.12), suppose $x_k(n) = x(n)$, i.e., all sensors observe an identical signal and furthermore, assume $x(n)$ is itself Gaussian with variance σ_s^2 . Thus the HT problem becomes a test between two multivariate Gaussian distributions

$$\mathcal{H}_0 : \mathbf{y}(n) \sim \mathcal{N}(0, \sigma_w^2 \mathbf{I}),$$

$$\mathcal{H}_1 : \mathbf{y}(n) \sim \mathcal{N}(0, \mathbf{C}),$$

where

$$\mathbf{C} = \begin{bmatrix} \sigma_s^2 + \sigma_w^2 & \sigma_s^2 & \cdots & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 + \sigma_w^2 & \cdots & \sigma_s^2 \\ \vdots & \ddots & \ddots & \sigma_s^2 \\ \sigma_s^2 & \sigma_s^2 & \cdots & \sigma_s^2 + \sigma_w^2 \end{bmatrix}.$$

Notice that this example is a slight generalization for the example described in [35, Section V.C] which deals with real variables. The proof of optimality of a threshold test of $|y(n)|^2$ carries over to the complex case. However, the result applies only to the case where $N = 1$. Notice that here the optimality is weaker than that for a single node with the same signal model; there, the energy detector is optimal for all N .

3.2.2 PSK Signal over Fast Rayleigh Fading Channel

Consider the model in (3.12) with $x_k(n) = h_k(n)s(n)$ and $h_k(n)$ is a complex Gaussian variable and is independent in k and n , $s(n)$ is drawn from a PSK constellation set \mathcal{S} defined in (3.5). From Section 3.1, we know that $x_k(n)$ is itself a complex Gaussian random variable. We show in the following that these $x_k(n)$ are indeed independent of each other for different k . Or more generally,

Lemma 3.1. *For PSK signals over Rayleigh fading channels, if the PSK sequence from the primary user is independent in time and the fading channels are independent across both k (sensors) and n (time), then the received signal vectors $\mathbf{y}_k = [y_k(1), \dots, y_k(N)]$ are independent across different sensors given either hypothesis.*

Proof. Under \mathcal{H}_0 , the independence is trivial since the noise variables are assumed independent across both time n and sensor k . Under \mathcal{H}_1 , $y_k(n) = h_k(n)s(n) + w_k(n)$. However, it was shown in Section 3.1 that $x_k(n) = h_k(n)s(n)$ is complex Gaussian distributed. We now show that for any $k \neq k'$, $x_k(n)$ and $x_{k'}(n')$ are independent of each other for any n and n' . For $n \neq n'$, the independence is trivially true as $(h_k(n), s(n))$ and $(h_{k'}(n'), s(n'))$ are independent of each other. For $n = n'$, we show that $x_k(n)$ and $x_{k'}(n)$ are independent Gaussian distributed random variables.¹

Following similar approach in establishing the Gaussian distribution of $x_k(n)$, we have

$$\begin{aligned}
 p(x_k, x_{k'}) &= \sum_{m=1}^M p(x_k, x_{k'}, s_m) \\
 &= \sum_{m=1}^M p(x_k, x_{k'} | s_m) \pi_m \\
 &\stackrel{(a)}{=} \sum_{m=1}^M p(x_k | s_m) p(x_{k'} | s_m) \pi_m \\
 &\stackrel{(b)}{=} p_{h_k}(x_k) p_{h_{k'}}(x_{k'})
 \end{aligned}$$

¹Notice that proving uncorrelatedness between $x_k(n)$ and $x_{k'}(n)$ is not sufficient; uncorrelated Gaussian random variable does not necessarily imply independence unless the two are jointly Gaussian.

where (a) follows as $x_k(n)$ and $x_{k'}(n)$ are conditionally independent given $s(n)$ and (b) is because given $s(n) = s_m$, x_k and $x_{k'}$ respectively have the exact distribution of $h_k(n)$ and $h_{k'}(n)$. Thus we have shown that $x_k(n)$ and $x_{k'}(n)$ are independent Gaussian random variables. \square

Now that we have proved that the observations y_k are conditionally independent across k given either hypothesis, local LRT at each sensor is optimal [4, 34]. Given that the LRT is equivalent to an energy detector for PSK in independent Rayleigh fading channels, as was established in Section 3.1, the energy detector is therefore also optimal for cooperative spectrum sensing.

3.2.3 QAM Signal over Fast Rayleigh Fading Channel

For QAM signal in Rayleigh fading channels, it was already established that the energy detector is not optimal in the single node case for the general case of $N > 1$. Thus for the distributed case, we only consider the simple case of $N = 1$, i.e., using a single observation. From (3.8), it is clear that given \mathcal{H}_1 , $y_k(n)$'s are not independent for different k .

We now use the approach in [35] to show that a threshold test of $|y_k(n)|^2$ is still optimal for $N = 1$. As we only consider one sample, in the following we suppress the time index n . Let us first define a hidden variable such that the observations become independent given the hidden variable. Define

$$W = |s|I(\mathcal{H} = \mathcal{H}_1) = \begin{cases} 0 & \mathcal{H} = \mathcal{H}_0 \\ |s| & \mathcal{H} = \mathcal{H}_1, \end{cases}$$

where $I(\mathcal{H} = \mathcal{H}_1)$ is an indicator function. It is clear that $\mathcal{H} - W - \mathbf{y}$ forms a Markov chain, i.e., given W , the distribution of \mathbf{y} is independent of \mathcal{H} . To see this, if $W = 0$, then \mathbf{y} is the noise vector \mathbf{w} , thus is independent of \mathcal{H} . Conditioned on $W = |s_m|$, we know that this is equivalent to sending a PSK symbol with signal energy $|s_m|^2$. The distribution \mathbf{y} is again

dependent only on $|s_m|$ given the circular invariance of complex Gaussian. Additionally, W induces conditional independence of \mathbf{y} , again, for the same reason as using PSK, the observations are conditionally independent. Thus W serves as the hidden variable in the framework described in [35]. However, Proposition 3.1 does not directly apply here given that the observations are complex (i.e., the monotone property defined for $y_k > y_{k'}$ is not applicable).

Nevertheless, an alternative proof can be constructed. First, given $W = |s|$, y_k is a circularly invariant complex Gaussian random variable. As such, $W - |y_k| - y_k$ forms a Markov chain, i.e., given $|y_k|$, y_k is independent of W . Given that y_k 's are conditionally independent given W , $(|y_1|, \dots, |y_K|)$ forms a sufficient statistic for W [36]. Thus they also form a sufficient statistic for \mathcal{H} . Now that we can instead consider $|y_k|$'s as local observations, it is then a straightforward exercise to check that the conditions in Proposition 3.1 is satisfied with the defined W . As such, a threshold test of $|y_k|$, or equivalently $|y_k|^2$ is optimal for the cooperative sensing problem with QAM in Rayleigh fading channels and $N = 1$.

3.3 Summary

This chapter studied the optimality of an energy detector for spectrum sensing for both a stand-alone system and a cooperative spectrum sensing system. For the single node case, it was shown that for all the cases considered in this chapter, the energy detector is either provably optimal or nearly optimal compared with the true optimum test. For cooperative spectrum sensing, however, optimality was established only for several special cases. These include PSK over independent Rayleigh fading for arbitrary sample numbers, and the special case of Gaussian signals in Gaussian channels and QAM signals in Rayleigh fading channels when the detector is limited to using a single sample.

CHAPTER 4

DECENTRALIZED ESTIMATION WITH CORRELATED OBSERVATIONS IN A TANDEM NETWORK

Consider the simplest estimation model with two distributed sensors collecting noisy observations of a parameter θ :

$$X = \theta + W_1, \quad Y = \theta + W_2, \quad (4.1)$$

where $(W_1, W_2) \sim \mathcal{N}(0, 0, \sigma_1^2, \sigma_2^2, \rho)$, i.e., the noises are bivariate normal distributed with zero mean, respective variances σ_1^2, σ_2^2 , and correlation coefficient ρ . Without loss of generality, $\sigma_1^2 \geq \sigma_2^2$ is assumed throughout this chapter unless otherwise stated. Of particular interest to the present chapter is a tandem network, where one node serves as the FC, and makes decisions based on its own observation as well as the compressed (quantized) input received from the other node. Such a two-node tandem network is illustrated in Fig. 4.1. With Gaussian model, a natural criterion for evaluating the estimation performance is the mean squared error (MSE). With distributed estimation and quantized observations, com-

puting MSE is often cumbersome and even intractable. Instead, Cramer-Rao lower bound (CRLB) [37, 38] is often used which is equivalent to evaluating the Fisher information (FI) given the observation model. Thus we will primarily use FI in assessing distributed estimators with dependent observations; as CRLB is not always tight with quantized observations, we will also resort to MSE evaluation when feasible.

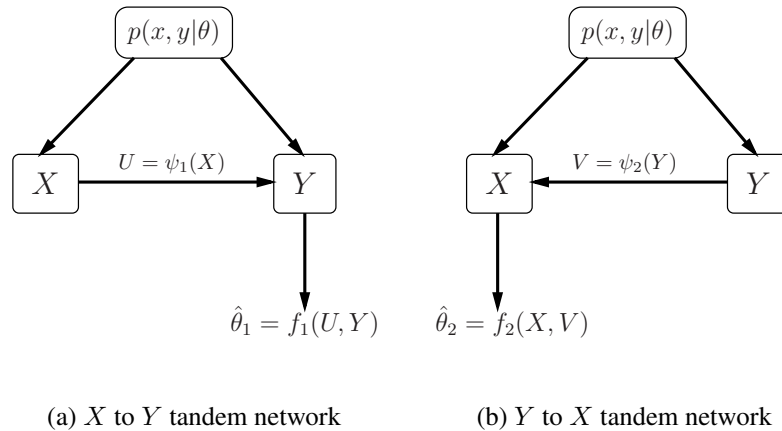


Fig. 4.1: Two fusion structures under a communication constraint where $\psi_i(\cdot)$ is assumed to be a one-bit quantizer in the present chapter

Local quantizer design in a distributed estimation system has been examined earlier in, e.g., [5] [39], where the focus is largely on algorithmic design and the accompanied numerical results. More concrete analytical results have been obtained in [40] [41], where score-function quantizer (SFQ) is shown to be optimal for maximizing the Fisher information (FI). The result, however, is derived under the assumption that observations are conditionally independent.

In the absence of the CI assumption, decentralized inference becomes much more challenging [42, 43]. In this case, the *structure* of the optimal quantizer at local sensors is usually coupled with other nodes. This difficulty is much well understood for *distributed detection* with dependent observations. Under the CI assumption, likelihood-ratio quantizers (LRQ) at local sensors have been shown to be optimal under both the Bayesian and Neyman-Pearson frameworks [44]. They are also known to maximize the Kullback-Leibler distance and Chernoff information for distributed detection, again, under the CI

assumption [45]. With this assumption removed, LRQs at local sensors are often not optimal [46,47], and the quantizer design becomes NP hard for the general dependent case [45]. The problem of interest to the present chapter, namely decentralized estimation with dependent observations has yet to receive much attention. An early work [39] presented some numerical results of threshold quantizer design with dependent observations, though no analytical result has been obtained.

By obtaining and understanding the optimum quantizer design for decentralized estimation with dependent observations, we are able to answer some of the interesting and important questions arising in decentralized inference. We address two of them in this chapter: 1) what is the preferred communication direction in tandem networks so that the ultimate inference performance at the FC is optimized, and 2) how does data dependency influence the inference performance compared with that of the independent case?

The communication direction problem can be illustrated using the simple two-node tandem network. Figs. 4.1(a) and (b) represent two different configurations where either X or Y serves as the FC; the question is which one yields better performance given a joint distribution $p(x, y)$. This issue was first introduced and analyzed in distributed detection with CI assumption under either hypothesis [48]. In this early work, the optimal configuration is proved to be dependent on external factors, such as prior probabilities of hypotheses and cost assignments. The authors in [49] addressed a similar problem under the additive correlated Gaussian noise model; by restricting the peripheral nodes to implement LRQ, the authors established that the preferred communication direction depends on the correlation across sensor observations. More recently, for the same model considered in [49], the optimal decision structure is obtained for the peripheral node in [50], and it was shown that with additive Gaussian noises, having the better sensor (i.e., the one with higher SNR) serving as the FC is always preferred regardless of the correlation coefficient.

The ultimate goal in this section is to measure the impact of data correlation on the inference performance. As mentioned before, data dependency is usually believed to bring

in data redundancy as it reduces the effective sample size [33, Ch. 5]. However, there are counter-examples showing that data correlation can be exploited to significantly benefit the inference performance, such as the noise canceling effect in the case of negatively correlated noises in centralized estimation problems [51, Ch. 6]. Compared to the centralized setup, the problem of how data correlation affects estimation in decentralized systems is much more complicated, and it is one of the problems of interest in this chapter.

4.1 Fisher Information

4.1.1 Classical Setting

In an inference problem, an observation X carries information about an unknown parameter θ , which the probability of X depends on, and that amount of information is characterized by FI. The likelihood function of θ is the probability of X conditioned on the value of θ , and the score function, denoted as $S_\theta(x)$ is defined as the partial derivative of the log-likelihood function with respect to θ ,

$$S_\theta(x) = \frac{\partial}{\partial \theta} \log p(x|\theta).$$

In classical statistics, the FI is defined as the expectation of the second moment of the score,

$$\begin{aligned} J(X; \theta) &= E \left[\left(\frac{\partial}{\partial \theta} \log p(x|\theta) \right)^2 \right] \\ &= E [S_\theta^2(x)]. \end{aligned}$$

The inverse of FI, $J(X; \theta)^{-1}$, namely CRLB, serves as a lower bound on the variance of estimators of a deterministic parameter. Particularly, for any unbiased estimator, the estimation MSE is at least as high as the FI, and the one which achieves the CRLB is an efficient estimator. Nonetheless, there may not exist an unbiased estimator which attains the bound. The CRLB can be also used to bound the variance of biased estimators given

the bias, and it is possible that some biased estimators yield variance and MSE lower than the CRLB.

4.1.2 Bayesian Framework

In the Bayesian framework, the unknown parameter θ is treated as a random variable. Therefore, the information carried by the observation X about the parameter can not be subject to a particular value of θ , and the parameter itself carries a certain amount of prior information which has to be incorporated. The Bayesian Fisher information ¹ is defined as [37]

$$\begin{aligned} J_B(X; \theta) &= E_{X, \Theta} \left[\left(\frac{\partial}{\partial \theta} \log p(x, \theta) \right)^2 \right] \\ &\triangleq J_D(X; \theta) + J_P(\theta) \end{aligned}$$

if the expectation exists. The first term

$$J_D(X; \theta) = E_{\Theta} E_{X|\theta} \left[\left(\frac{\partial}{\partial \theta} \log p(x|\theta) \right)^2 \right]$$

is the FI associated with data averaged over θ , and the second term

$$J_P(\theta) = E_{\Theta} \left[\left(\frac{\partial}{\partial \theta} \log p(\theta) \right)^2 \right]$$

is the prior information associated with $p(\theta)$. Similar to the classical case, the PCRLB, defined as $J_B(X; \theta)^{-1}$, provides a theoretical lower bound for the estimation MSE of any estimator under the Bayesian framework.

As with other meaningfully defined information quantities, such as mutual information,

¹Note that there are different names for Fisher information when prior information is incorporated, e.g., Bayesian information in [37]. We will simply refer to them as Fisher information which takes a form that incorporate prior distribution whenever applicable.

the averaged FI satisfies the chain rule [52], i.e.,

$$J_D(X, Y; \theta) = J_D(Y; \theta) + J_D(X; \theta|Y),$$

where the conditional FI is defined as

$$J_D(X; \theta|Y) = E_{\Theta} E_{X, Y|\theta} \left[\left(\frac{\partial}{\partial \theta} \log p(x|y, \theta) \right)^2 \right].$$

Therefore, the joint Bayesian FI can be decomposed as

$$J_B(X, Y; \theta) = J_D(Y; \theta) + J_D(X; \theta|Y) + J_P(\theta). \quad (4.2)$$

Notice that with independent Gaussian noises in (4.1), the joint likelihood function factorizes into the product of the marginal distributions, leading to

$$J_B(X, Y; \theta) = J_D(Y; \theta) + J_D(X; \theta) + J_P(\theta). \quad (4.3)$$

Clearly, from (4.2) and (4.3), the effect of data correlation on the estimation performance amounts to comparing $J_D(X; \theta|Y)$ and $J_D(X; \theta)$, provided that there exists an estimator that attains the PCRLB. Redundancy occurs when $J_D(X; \theta|Y) < J_D(X; \theta)$, i.e., FI decreases with conditioning. For the sake of simplicity, the subscript of $J_D(\cdot)$ is dropped without any ambiguity.

4.2 Centralized Estimation under Gaussian Noise

Consider the estimation problem in the bivariate Gaussian model described in (4.1). It is straightforward to show that, with $|\rho| < 1$

$$J(X, Y; \theta) = \frac{1}{1 - \rho^2} \cdot \left[\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} - \frac{2\rho}{\sigma_1\sigma_2} \right], \quad (4.4)$$

$$J(X; \theta) = \frac{1}{\sigma_1^2}, \quad (4.5)$$

$$J(Y; \theta) = \frac{1}{\sigma_2^2}. \quad (4.6)$$

For the case with $|\rho| = 1$, the observations are perfectly correlated, indicating the degenerate cases of bivariate Gaussian, so the corresponding FI can be obtained as taking the appropriate limit of $J(X, Y; \theta)$ when $\rho \rightarrow \pm 1$, where the limit is defined in the usual sense.

To examine the effect of data correlation, we now compare $J(X; \theta)$ and $J(X; \theta|Y)$ for the Gaussian additive model. From (4.4), (4.5), and (4.2), we get

$$J(X; \theta|Y) = \frac{1}{1 - \rho^2} \cdot \left(\frac{\rho}{\sigma_2} - \frac{1}{\sigma_1} \right)^2. \quad (4.7)$$

Comparing (4.7) with (4.5), we can categorize various correlation regimes in terms of inference performance relative to that of the independence case.

- $\rho = -1$. As X and Y are perfectly and negatively correlated (i.e., $\sigma_2 W_1 = -\sigma_1 W_2$), complete noise cancellation can be achieved by linearly combining X and Y . Notice that in the additive Gaussian model, the optimal (MMSE) estimator happens to be linear. The optimal estimator for θ is

$$\hat{\theta} = \frac{\sigma_2}{\sigma_1 + \sigma_2} x + \frac{\sigma_1}{\sigma_1 + \sigma_2} y,$$

which can be shown to be identically θ hence the MSE is 0. Equivalently, we obtain

the limiting FI when $\rho \rightarrow -1$,

$$J(X; \theta|Y) = \lim_{\rho \rightarrow -1} \frac{\left(\frac{1}{\sigma_1} - \frac{\rho}{\sigma_2}\right)^2}{1 - \rho^2} \rightarrow \infty,$$

which is consistent with the fact that perfect estimation can be achieved with $\rho = -1$.

- $-1 < \rho < 0$. For this regime, it is straightforward to show that partial noise cancellation is optimal in the sense of minimizing the MSE:

$$\hat{\theta} = \frac{\sigma_2^2 - \rho\sigma_1\sigma_2}{\sigma_2^2 + \sigma_1^2 - 2\rho\sigma_1\sigma_2}x + \frac{\sigma_1^2 - \rho\sigma_1\sigma_2}{\sigma_2^2 + \sigma_1^2 - 2\rho\sigma_1\sigma_2}y$$

with the corresponding MSE $\frac{(1-\rho^2)\cdot\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}$. Equivalently, the FI for this case can be shown to be greater than that of the independent case, i.e.,

$$J(X; \theta|Y) - J(X; \theta) = \frac{\frac{\rho^2}{\sigma_1^2} - \frac{2\rho}{\sigma_1\sigma_2} + \frac{\rho^2}{\sigma_2^2}}{1 - \rho^2} > 0.$$

Thus, negatively correlated Gaussian noises in the additive model (4.1) is always beneficial for the estimation performance. Notice that in the centralized case Gaussian model, the minimum mean squared error (MMSE) estimator - coincides with the linear minimum mean squared error (LMMSE) estimator. This is not the case with the decentralized case where the peripheral node needs to quantize its observation.

- $0 \leq \rho \leq \frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}$. This is when dependency implies redundancy, i.e.,

$$J(X; \theta|Y) - J(X; \theta) = \frac{\rho^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) - \frac{2\rho}{\sigma_1\sigma_2}}{1 - \rho^2} \leq 0.$$

Thus data dependence negatively affects the estimation performance compared with that of the independent case. At the particular point $\rho = \frac{\sigma_2}{\sigma_1}$, the conditional FI becomes 0, implying that X is completely redundant given Y . This is the consequence

of the following Markov chain $\theta - Y - X$ when $\rho = \frac{\sigma_2}{\sigma_1}$, i.e., X is independent of θ given Y .

- $\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} < \rho < 1$. This is the parameter regime where positively correlated noises also benefit the inference performance. Checking the FI, it is easy to show that for this parameter regime,

$$J(X; \theta|Y) - J(X; \theta) = \frac{\rho^2 \left(\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} \right) - \frac{2\rho}{\sigma_1\sigma_2}}{1 - \rho^2} > 0.$$

To understand why this is the case, we note that for positively correlated noises, (partial) noise cancellation is also attainable by subtracting one observation from the other with proper scaling. However, the subtraction also reduces the signal power. The balancing point happens to be at

$$\rho = \frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2}, \quad (4.8)$$

i.e., beyond this value, noise cancellation more than compensates for the signal power reduction, resulting in improved estimation performance.

- $\rho = 1$. This is the extreme case when $\sigma_2 W_1 = \sigma_1 W_2$. Depending on the values of σ_1 and σ_2 , there are two distinct cases that have completely different ramifications on the underlying estimation problem.

– $\sigma_1 = \sigma_2$, i.e., $W_1 = W_2$, thus $X = Y$. Therefore one of the observations is completely redundant. The conditional FI is now

$$J(X; \theta|Y) = \frac{1}{\sigma_1^2} \lim_{\rho \rightarrow 1} \frac{(1 - \rho)^2}{1 - \rho^2} = 0.$$

Indeed, this is exactly a special case of the extreme point corresponding to $\rho = \frac{\sigma_2}{\sigma_1}$ hence X is redundant given Y . In this case, $\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} = 1$ hence the

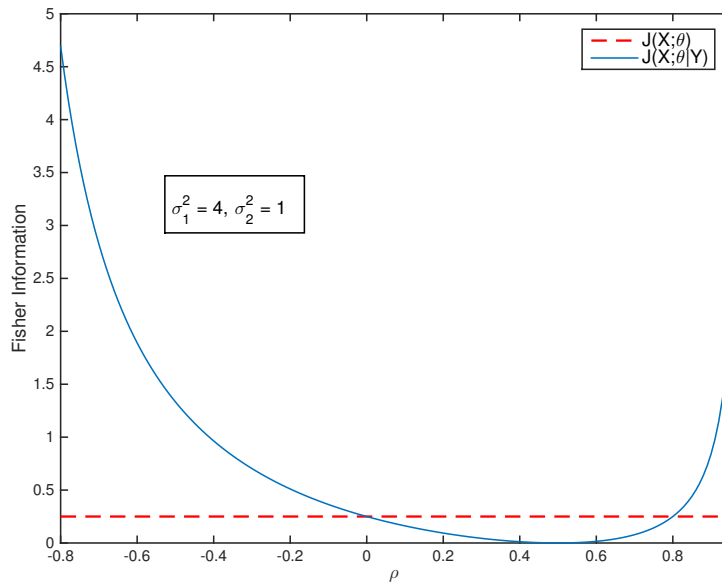


Fig. 4.2: Comparison between $J(X; \theta)$ and $J(X; \theta|Y)$

previous parameter regime $\frac{2\sigma_1\sigma_2}{\sigma_1^2 + \sigma_2^2} < \rho < 1$ vanishes.

- $\sigma_1 > \sigma_2$. In this case, perfect estimation is achieved by the following estimator

$$\hat{\theta} = \frac{\sigma_1}{\sigma_1 - \sigma_2} y - \frac{\sigma_2}{\sigma_1 - \sigma_2} x.$$

Not surprisingly, the corresponding FI becomes unbounded as that when $\rho = -1$ and perfect estimation is also achieved.

$$J(X; \theta|Y) = \lim_{\rho \rightarrow -1} \frac{\left(\frac{1}{\sigma_1} - \frac{1}{\sigma_2}\right)^2}{1 - \rho^2} \rightarrow \infty.$$

The FI of the various parameter regimes are plotted in Fig. 4.2 where the noise variances are fixed as $\sigma_1 = 2$ and $\sigma_2 = 1$. With this setting, we can compute the boundary point (4.8) to be $\rho = 0.8$. From the figure, it is clear that $J(X; \theta|Y) \geq J(X; \theta)$ for $\rho \in [-1, 0]$ and $\rho \in [0.8, 1]$, while the opposite inequality holds for $\rho \in (0, 0.8)$. Additionally, the conditional FI $J(X; \theta|Y) = 0$ at $\rho = \frac{1}{2} = \frac{\sigma_2}{\sigma_1}$ indicating that X is completely redundant given Y .

4.3 Decentralized Estimation with Bivariate Gaussian Noises

We now proceed to consider the problem when X or Y is subject to quantization prior to being available at the other node that estimates θ . The requirement is often times due to various system constraints which collectively impose a finite capacity constraint for the communication between the two nodes. For either of the two configurations in Fig. 4.1, we first address the optimal quantizer design at the remote node. Subsequently, we attempt to answer the question of which node should be used as a FC for better estimation performance. Finally, the resulting quantizer structure as well as the optimal communication direction will reveal how correlation may impact the estimation performance.

4.3.1 The Optimality of Single Threshold Quantizer

For simplicity, we assume the extreme case of a one-bit quantizer at the local sensor. Consider Fig. 4.1(a). The FI at the estimator decomposes into three terms

$$J(Y, U(X); \theta) = J(Y; \theta) + J(U(X); \theta|Y) + J(\theta),$$

where $U(X)$ is the binary quantizer output for X . Hence maximizing the overall FI $J(Y, U(X); \theta)$ is equivalent to maximizing the conditional term $J(U(X); \theta|Y)$. A special case is when the noises are independent. In this case, the second term becomes independent of Y , i.e., one only need to design a quantizer at X such that the FI of the quantizer output is maximized. The dependent case is more complicated since Y is not accessible at node X , therefore it is not realistic to design a quantizer to maximize FI for each specific value of Y . However, we will show later that the *class* of optimal quantizer is in fact independent of Y .

The definition of the score-function quantizer (SFQ) is given in [40].

Definition 4.1. A quantizer $\psi(\cdot)$ is a monotone SFQ with threshold vector $t = (t_1, \dots, t_{D-1}) \in \mathbf{R}^{D-1}$, if $\psi(x) = d \Leftrightarrow S_\theta(x) \in [t_{d-1}, t_d]$ for any x , where $t_0 = -\infty \leq t_1 \leq \dots \leq t_{D-1} \leq \infty = t_D$. Any permutation of the monotone SFQ generated by a permutation mapping $\pi : \{1, \dots, D\} \rightarrow \{1, \dots, D\}$ is a SFQ.

The significance of SFQ is that it optimizes the FI $J(U(X); \theta)$ among all quantizers with the same quantization level. In a tandem network, to maximize the conditional FI $J(U(X); \theta|Y)$, the corresponding notion of the score-function is characterized by the conditional distribution as

$$S_\theta(x|y) = \frac{\partial}{\partial \theta} \log p(x|\theta, y), \quad (4.9)$$

which is dependent on both θ and y . Therefore, we denote the SFQ corresponding to $S_\theta(x|y)$ as $\psi_{\theta,y}(x)$ for the time being. The reason that SFQ $\psi_{\theta,y}(x)$ can not directly apply to our case is two-fold: 1) Score-function is dependent on the parameter θ , whose value is unknown *a priori* and it is generally not possible to design a single quantizer that is optimal for every θ [41]; and 2) in a tandem network, the observation from the other node (Y) is not available at the quantizer node (X). However, by defining the class of SFQ's at a particular pair (θ, y) as $\Psi_{\theta,y} = \{\psi_{\theta,y}(x) : \psi_{\theta,y}(x) \text{ is a SFQ at } (\theta, y)\}$, and then by using a similar argument as in [41], it can be shown that if the conditional distribution $p(x|\theta, y)$ satisfies a monotonicity property, then the class of SFQ's at every single pair of (θ, y) is identical. This is summarized in the following lemma.

Lemma 4.1. Let $T(X)$ be a function of X . If the score function can be expressed as $S_\theta(x|y) = f_{\theta,y}(T(x))$, where $f_{\theta,y}(\cdot)$ is monotone increasing for any (θ, y) , then

1. the class of SFQ's, $\Psi_{\theta,y}$, is identical for all (θ, y) pairs, i.e., for any (θ, y) and (θ', y') ,

$$\Psi_{\theta,y} = \Psi_{\theta',y'};$$
2. every SFQ is equivalent to a quantizer on $T(x)$ with $D - 1$ thresholds, i.e. there exist

$-\infty = t'_0 \leq t'_1 \leq \dots \leq t'_{D-1} \leq t'_D = \infty$ such that

$$\psi(x) = d \Leftrightarrow T(x) \in [t'_{d-1}, t'_d].$$

Proof. Since $S_\theta(x|y)$ is monotone increasing in $T(x)$, any SFQ is equivalent to quantizing $T(x)$ while retaining the order of thresholds. Therefore, the class of SFQ's is independent of θ and y . \square

For the problem of a tandem network with bivariate Gaussian noises, $S_\theta(x|y)$ is determined by

$$\begin{aligned} S_\theta(x|y) &= \frac{\partial}{\partial \theta} \log \frac{1}{\sqrt{2\pi(1-\rho^2)\sigma_1^2}} e^{-\frac{(x-\theta-\frac{\sigma_1}{\sigma_2}\rho(y-\theta))^2}{2(1-\rho^2)\sigma_1^2}} \\ &= \frac{\partial}{\partial \theta} \left[-\frac{\left(x-\theta-\frac{\sigma_1}{\sigma_2}\rho(y-\theta)\right)^2}{2(1-\rho^2)\sigma_1^2} \right] \\ &= \frac{1-\frac{\sigma_1}{\sigma_2}\rho}{(1-\rho^2)\sigma_1^2} \left(x - \frac{\sigma_1}{\sigma_2}\rho y - \left(1-\frac{\sigma_1}{\sigma_2}\rho\right)\theta \right), \end{aligned}$$

It is clear that we can choose $T(x) = x$ for $-1 \leq \rho < \frac{\sigma_2}{\sigma_1}$, and $T(x) = -x$ for $\frac{\sigma_2}{\sigma_1} < \rho \leq 1$, such that Lemma 4.1 applies. Therefore, we can straightforwardly show that the optimal one-bit quantizer in the Gaussian additive model is a single threshold quantizer on the observation itself, i.e.

$$\psi(X) = \mathbf{1}\{x \geq \gamma\}.$$

Remark: *Conditional sufficient statistics can help to find the function $T(X)$. We give the definition as follows.*

Definition 4.2. *A statistic $W(X)$ is a conditional sufficient statistic for θ conditioned on Y , if the conditional distribution of the sample X given the value of $W(X)$ and Y does not depend on θ . [53].*

The Neyman-Fisher factorization theorem can be generalized to characterize conditional sufficient statistics: $W(X)$ is a conditional statistic for θ given Y iff $p(x, y|\theta) = g_\theta(W(x), y) \cdot h(x, y)$ for all x, y and θ [53]. In this case, the score function is also a function of $W(X)$. Then if the score function is monotone on $W(X)$, one can simply set $T(X) = W(X)$.

The optimal threshold $\gamma^* \in \mathbf{R}$ is one such that

$$\gamma^* = \arg \max_{\gamma} J(\psi(X, \gamma); \theta|Y),$$

where we use the notation $\psi(X, \gamma)$ to indicate the dependence of the quantizer output on the threshold γ . By definition, the conditional FI is characterized as

$$\begin{aligned} & J(\psi(X, \gamma); \theta|Y) \\ &= E_{X,Y,\Theta} [S_\theta(\psi(X, \gamma)|Y)^2] \\ &= \int_{\Theta} f(\theta) \int_{\mathcal{Y}} g(y|\theta) \sum_{u \in \{0,1\}} P(\psi(x, \gamma) = u|\theta, y) S_\theta(\psi(x, \gamma)|y)^2 dy d\theta, \end{aligned} \quad (4.10)$$

where $f(\theta)$ is the pdf of θ , $f(y|\theta)$ is the conditional distribution of Y given θ , and $S_\theta(\psi(x, \gamma)|y)$ is defined according to (4.9).

We provide the following theorem to identify the optimal threshold γ^* for the one-bit quantizer in the additive Gaussian model.

Theorem 4.1. *Let $\theta \sim \mathcal{N}(\mu, \sigma^2)$ be a parameter with Gaussian distribution. Let X and Y be noisy observations of θ with bivariate additive Gaussian noise. Then, for the tandem system in Fig. 4.1(a), the optimal one-bit quantizer on X is a single threshold quantizer on X , and the optimal threshold is $\gamma^* = \mu$, i.e., $U(X) = \psi(X, \mu)$.*

Without loss of generality, we assume $\mu = 0$. We begin with the following lemma to prove Theorem 4.1.

Lemma 4.2. Define the function $\eta(t) = \frac{e^{-t^2}}{Q(t)(1-Q(t))}$. Then $\eta(t)$ is a symmetric function with respect to $t = 0$, while monotonically decreasing in $t \in (0, \infty)$, and

$$t^* = \arg \max_t \eta(t) = 0, \quad \eta^* = 4.$$

Proof. The detailed proof is given in Appendix C. □

With the Gaussian assumption described in Theorem 4.1, conditional FI in (4.10) can be expanded into

$$\begin{aligned} J(\psi(X, \gamma); \theta|Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{e^{-\frac{\theta^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \cdot \frac{e^{-\frac{(y-\theta)^2}{2\sigma_2^2}}}{\sqrt{2\pi\sigma_2^2}} \\ &\quad \cdot \frac{\frac{(\rho\sigma_1 - \sigma_2)^2}{2\pi(1-\rho^2)\sigma_1^2\sigma_2^2} \cdot e^{-\left(\frac{\gamma - \theta - \frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right)^2}}{Q\left(\frac{\gamma - \theta - \frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right) \left(1 - Q\left(\frac{\gamma - \theta - \frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right)\right)} dy d\theta, \end{aligned} \quad (4.11)$$

Observe that the integrand can be made symmetric for (y, θ) around $(0, 0)$ by setting $\gamma = 0$ (c.f. Lemma 4.2). Then this intuitively attains the maximum and we give a formal proof below.

Proof. With the assumption $\mu = 0$, we only need to show

$$J(\psi(X, 0); \theta|Y) - J(\psi(X, \gamma); \theta|Y) > 0, \quad \forall \gamma \neq 0. \quad (4.12)$$

We first rewrite the integral in (4.11) as

$$J(\psi(X, \gamma); \theta|Y) \stackrel{w=y-\theta}{\propto} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\theta)g(w)\eta(\beta\gamma - \beta\theta - \alpha w) dw d\theta,$$

where we denote $\alpha = \frac{\rho}{\sqrt{1-\rho^2}\sigma_2}$, and $\beta = \frac{1}{\sqrt{1-\rho^2}\sigma_1}$, and $f(\cdot)$ and $g(\cdot)$ as the pdf's of θ and

$Y|\theta$, respectively. Plug it back into the left-hand side of (4.12), we have

$$\begin{aligned}
& \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\theta)g(w)\eta(-\beta\theta - \alpha w) d\theta dw - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\theta)g(w)\eta(\beta\gamma - \beta\theta - \alpha w) d\theta dw \\
\stackrel{(1)}{=} & \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\theta)g(w)\eta(\beta\theta + \alpha w) d\theta dw - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\theta)g(w)\eta(\beta\gamma - \beta\theta - \alpha w) d\theta dw \\
\stackrel{(2)}{=} & \int_{-\infty}^{\infty} \int_{-\infty}^{\frac{\gamma}{2} - \frac{\alpha w}{\beta}} \left(f(\theta') - f\left(\gamma - \frac{2\alpha w}{\beta} - \theta'\right) \right) g(w) \\
& \times (\eta(\beta\theta' + \alpha w) - \eta(\beta\gamma - \beta\theta' - \alpha w)) d\theta' dw. \tag{4.13}
\end{aligned}$$

The first equality follows from Lemma 4.2, and the second one follows by splitting the integral with respect to θ at $\frac{\gamma}{2} - \frac{\alpha w}{\beta}$, and change of variable $\theta' = \gamma - \frac{2\alpha w}{\beta} - \theta$.

Without loss of generality, we first consider the interval of w such that $\frac{\gamma}{2} - \frac{\alpha w}{\beta} > 0$. Since both $f(\cdot)$ and $\eta(\cdot)$ are even functions symmetric at 0 and monotone decreasing for $t > 0$, we can infer that $f(\theta') - f(\gamma - \frac{2\alpha w}{\beta} - \theta')$ and $\eta(\beta\theta' + \alpha w) - \eta(\beta\gamma - \beta\theta' - \alpha w)$ are both non-negative, as $\theta' < \frac{\gamma}{2} - \frac{\alpha w}{\beta}$ in the inner integral; while for $\frac{\gamma}{2} - \frac{\alpha w}{\beta} < 0$, they are both non-positive. Moreover, since $g(w)$ is always positive, then the integrand in (4.13) is always non-negative. On the other hand, the integrand equals 0 in two cases: 1) $\gamma = 0$, and 2) $\frac{\gamma}{2} - \frac{\alpha w}{\beta} = 0$. For the first case, $\gamma = 0$ implies that $\eta(\beta\theta' + \alpha w) - \eta(\beta\gamma - \beta\theta' - \alpha w) = 0$ for any w , so the whole integral also equals 0. However, for the second case, the integrand is equal to 0 only for the particular point $w = \frac{\gamma\beta}{2\alpha}$, so the whole term integrates above 0 when $\gamma \neq 0$. Hence, (4.12) is proved, and thus the unique optimal threshold is $\gamma^* = 0$. \square

Intuitively, since both $f(\theta)$ and $g(w)$ are even functions around 0 and monotone decreasing as the corresponding parameter deviates further away from 0 (bell-shaped), it is clear that maximizing the integral in (4.13) would result in placing the function $\eta(\cdot)$ around the origin. It is also clear from the above proof that the Gaussian assumption on the parameter θ in Theorem 4.1 can be generalized to an arbitrary distribution with symmetric bell-shaped pdf.

The PCRLB at the FC with different correlation coefficients are given in Fig. 4.3. In

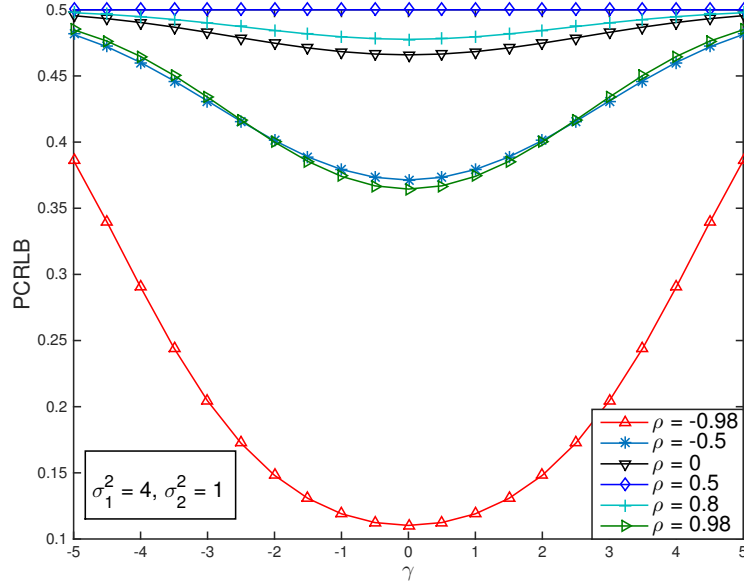


Fig. 4.3: PCRLB vs. γ for different ρ .

this figure, the parameter is standard normal distributed, and for different ρ , the threshold γ varies from -5 to 5 , which is a fairly large range compared to the signal and noise standard deviations. One can easily observe from the figure that the PCRLB is always minimized at $\gamma = 0$.

4.3.2 Communication Direction Problem

For a given joint distribution $p(x, y)$, the communication direction problem is equivalent to comparing the performance between the configurations in Fig. 4.1(a) and (b). Our goal is therefore to determine, for $\sigma_1^2 > \sigma_2^2$, which of the two FIs $J(U(X), Y; \theta)$ and $J(X, V(Y); \theta)$ is larger and whether the inequality depends on ρ , the correlation coefficient of the additive noises. Here $U(X)$ and $V(Y)$ are respectively the optimal quantizer outputs at X and Y , which are both single threshold quantizers at $\gamma^* = 0$.

For any arbitrary threshold $\gamma \in \mathbf{R}$, we can establish the following inequality,

Lemma 4.3. *Let $\psi(\cdot, \gamma)$ be a single threshold quantizer with threshold γ . For the observa-*

tions X and Y given in (4.1) and with $\sigma_1^2 > \sigma_2^2$, we have

$$J(\psi(X, \gamma), Y; \theta) > J(X, \psi(Y, \gamma); \theta), \quad \forall \gamma \in \mathbf{R} \quad (4.14)$$

Proof. See Appendix B. □

In the last section, we showed that the optimal thresholds for quantizing X and Y are both at $\gamma^* = 0$ for zero-mean Gaussian parameters. We show in the following that Lemma 4.3 implies that the preferred direction is independent of the prior distribution on θ . Indeed, it is true even if the parameter is treated as an unknown deterministic parameter.

Suppose γ^X and γ^Y represent the optimal thresholds for quantizing X and Y respectively, then it is trivial that

$$J(\psi(X, \gamma^X), Y; \theta) > J(\psi(X, \gamma^Y), Y; \theta).$$

On the other hand, Lemma 4.3 gives

$$J(\psi(X, \gamma^Y), Y; \theta) > J(X, \psi(Y, \gamma^Y); \theta).$$

The above two inequalities together imply that

$$J(U(X), Y; \theta) > J(X, V(Y); \theta), \quad \forall \rho,$$

where $U(X) = \psi(X, \gamma^X)$, and $V(Y) = \psi(Y, \gamma^Y)$.

The results on the preferred communication direction is summarized in the next theorem.

Theorem 1. *Let θ be a parameter in a tandem network under bivariate additive Gaussian noises. With single-bit quantizers, the best strategy is to always quantize the observation with worse noise, and choose the better one as the fusion center.*

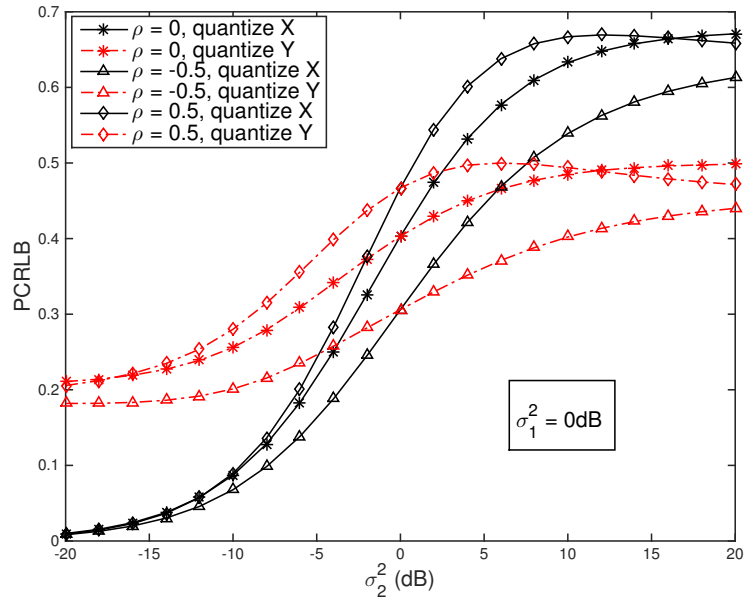


Fig. 4.4: PCRLB comparison with $\rho = -0.5, 0, 0.5$

We provide visualization of the PCRLB under different correlation coefficients in Fig. 4.4. In this figure, σ_1^2 , the noise variance at X , is fixed at 0dB, while σ_2^2 varies from -20 to 20 dB, and quantization is imposed on either X or Y . By comparing each pair of curves with the same ρ , one can clearly see that when $\sigma_2^2 < \sigma_1^2$, quantizing X yields smaller MSE, while the reverse is true with $\sigma_2^2 > \sigma_1^2$.

Notice that the curve of $\rho = 0.5$ and quantize Y increases and then starts to decrease at a particular point $\sigma_2^2 = 6$ dB. The reason is that in the case $\sigma_1 < \sigma_2$ and quantize Y , at the point $\rho = \frac{\sigma_1}{\sigma_2}$ ($\sigma_1^2 = 0$ dB and $\sigma_2^2 = 6$ dB implies that $\frac{\sigma_1}{\sigma_2} = 0.5 = \rho$), there forms a Markov chain $\theta - X - Y - V(Y)$, which implies that the quantized observation $V(Y)$ is totally redundant. It can be seen from (4.11) that it reduces to 0 when ρ takes that particular value.

4.3.3 Data Dependency and Redundancy

As illustrated in the centralized estimation example, data dependency could either imply redundancy or be exploited for better estimation. We now examine the same problem in a decentralized system.

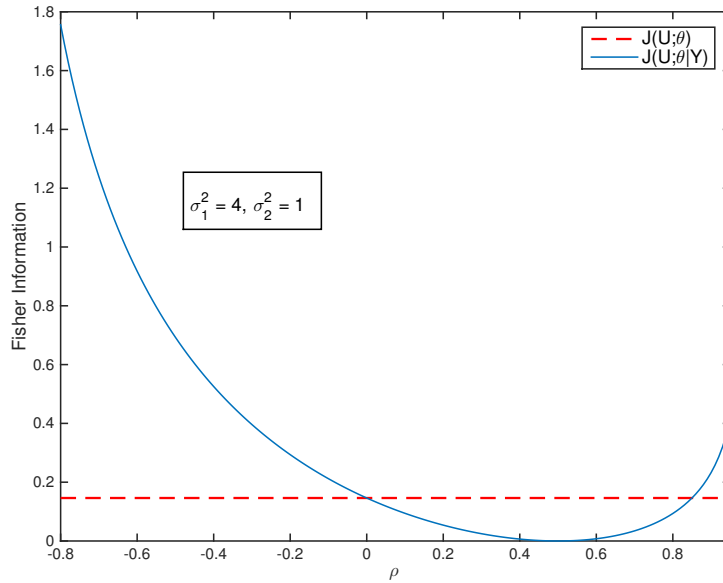


Fig. 4.5: Comparison between $J(U(X); \theta)$ and $J(U(X); \theta|Y)$

Similar to the centralized case, to evaluate data redundancy, we need to compare $J(U(X); \theta|Y)$ and $J(U(X); \theta)$, where

$$\begin{aligned}
 J(U(X); \theta|Y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{e^{-\frac{\theta^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \cdot \frac{e^{-\frac{(y-\theta)^2}{2\sigma_2^2}}}{\sqrt{2\pi\sigma_2^2}} \\
 &\quad \cdot \frac{\frac{(\rho\sigma_1 - \sigma_2)^2}{2\pi(1-\rho^2)\sigma_1^2\sigma_2^2} \cdot e^{-\left(\frac{-\theta - \frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right)^2}}{Q\left(\frac{-\theta - \frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right) \left(1 - Q\left(\frac{-\theta - \frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right)\right)} dy d\theta. \quad (4.15)
 \end{aligned}$$

and $J(U; \theta) = J(U; \theta|Y)|_{\rho=0}$. The numerical comparison is given in Fig. 4.5. As with the centralized case, negative correlation always benefits the estimation performance. For positive correlated noises, small correlation implies redundancy whereas large correlation can be exploited to enhance the estimation performance.

We now analytically establish the existence of different parameter regimes in terms of the correlation coefficient. As the closed-form expression for $J(U; \theta|Y)$ is intractable, we resort to a pair of upper and lower bounds.

Proposition 1. $J(U(X); \theta|Y)$ is upper bounded by

$$J(U(X); \theta|Y) < \frac{2(\sigma_2 - \rho\sigma_1)^2}{\pi(1 - \rho^2)\sigma_1^2\sigma_2^2} \triangleq J_u(\rho). \quad (4.16)$$

Proof. Directly apply Lemma 4.2. □

Proposition 2. $J(U(X); \theta|Y)$ is lower bounded by

$$J(U(X); \theta|Y) > \frac{2(\sigma_2 - \rho\sigma_1)^2\sigma_1}{\pi\sigma_1^2\sigma_2^2\sqrt{(1 - \rho^2)((1 + \rho^2)\sigma_1^2 + 2\sigma^2)}} \triangleq J_l(\rho), \quad (4.17)$$

Proof. Use the fact

$$\eta(t) = \frac{e^{-t^2}}{Q(t)(1 - Q(t))} \geq 4e^{-t^2}.$$

□

It can be verified that both $J_u(\rho)$ and $J_l(\rho)$ decrease monotonically from $\rho = -1$ to $\frac{\sigma_2}{\sigma_1}$, and increase from $\frac{\sigma_2}{\sigma_1}$ to $\rho = 1$, while achieving identical minimum 0 at $\frac{\sigma_2}{\sigma_1}$. Also, we notice that

$$\lim_{\rho \rightarrow \pm 1} J_u(\rho) = \lim_{\rho \rightarrow \pm 1} J_l(\rho) = \infty, \quad \sigma_1 > \sigma_2.$$

Hence, $J(U; \theta|Y)$ is also unbounded at the two extreme points. Since the bounds have finite values at $\rho = 0$, $J(U; \theta)$ is bounded, which means that there must be a boundary point ρ^* such that for any $\rho > \rho^*$, $J(U; \theta|Y) > J(U; \theta)$. Furthermore, it implies that negative correlation always yields higher FI than its independent counterpart, as does positive correlation beyond the boundary point. Between 0 and that boundary point, data correlation leads to redundancy, which negatively affects the estimation performance.

It is not surprising that at the two extreme points, the FI is unbounded. However, it does not necessarily imply perfect estimation at these extreme points. The reason is because the PCRLB is not tight for the case with quantized observations. To verify this, we compute the MMSE achieved by the conditional mean $E[\theta|u(x), y]$. In this case, the MMSE estimator

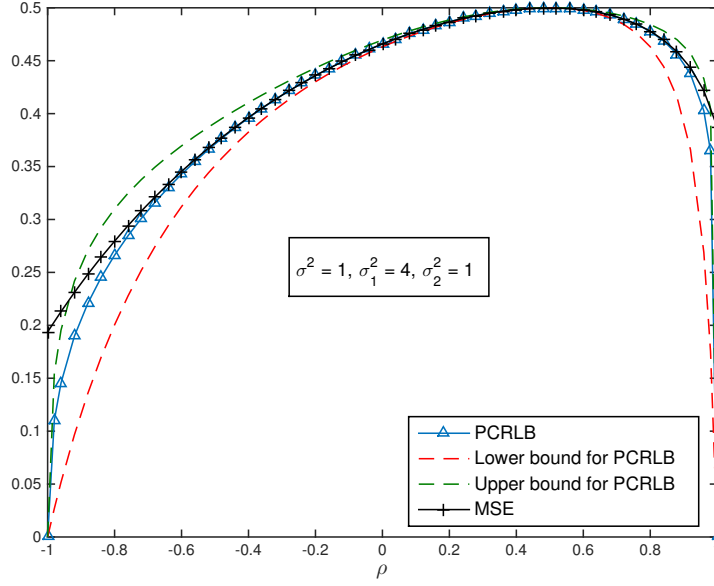


Fig. 4.6: CRLB, its bounds and MSE of the Decentralized Estimation Problem

is computed to be

$$\begin{aligned}\hat{\theta}(U = 0, y) &= \frac{\sigma^2}{\sigma^2 + \sigma_2^2}y - \frac{\sigma^2\sigma_2(\sigma_2 - \rho\sigma_1)}{\sqrt{2\pi(\sigma^2 + \sigma_2^2)^2D}} \frac{e^{-\frac{(\gamma - \kappa y)^2}{2D}}}{1 - Q\left(\frac{\gamma - \kappa y}{\sqrt{D}}\right)}, \\ \hat{\theta}(U = 1, y) &= \frac{\sigma^2}{\sigma^2 + \sigma_2^2}y + \frac{\sigma^2\sigma_2(\sigma_2 - \rho\sigma_1)}{\sqrt{2\pi(\sigma^2 + \sigma_2^2)^2D}} \frac{e^{-\frac{(\gamma - \kappa y)^2}{2D}}}{Q\left(\frac{\gamma - \kappa y}{\sqrt{D}}\right)},\end{aligned}\quad (4.18)$$

where $D = \frac{(1-\rho^2)\sigma_1^2\sigma_2^2 + \sigma^2\sigma_1^2 + \sigma^2\sigma_2^2 - 2\rho\sigma_1\sigma_2\sigma^2}{\sigma^2 + \sigma_2^2}$, and $\kappa = \frac{\sigma^2 + \rho\sigma_1\sigma_2}{\sigma^2 + \sigma_2^2}$. The first term $\frac{\sigma^2 y}{\sigma^2 + \sigma_2^2}$ is the optimal estimator by using only Y , and the second term can be viewed as a correction term that takes $U(X)$ into consideration.

In Fig. 4.6, we plot the numerically computed PCRLB, along with its lower and upper bounds derived in (4.16) and (4.17). Also plotted is the MMSE obtained using the above estimator. Clearly, the PCRLB is no longer tight in the correlation regimes close to the two extreme points, i.e., $\rho = \pm 1$. This is seen from the divergence between the MMSE and the numerically computed PCRLB.

Since in the decentralized estimation case, the PCRLB is not tight especially at the

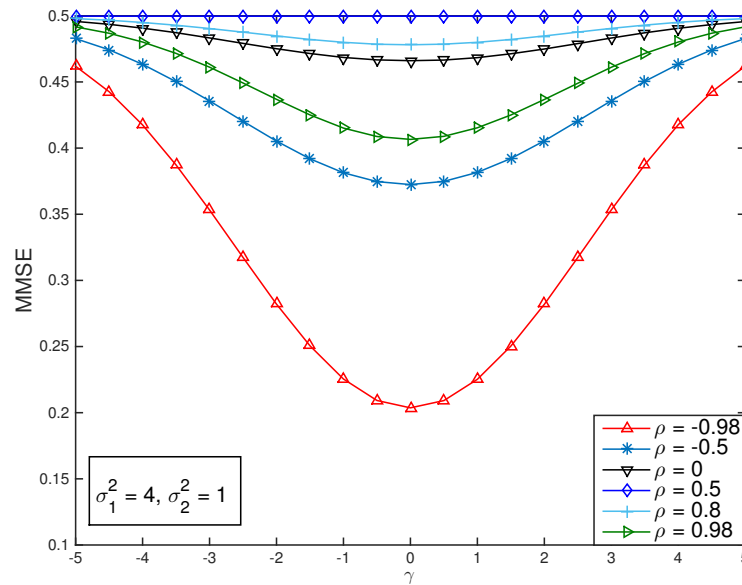


Fig. 4.7: MSE vs. γ for different ρ

extreme points, people may question that whether the optimal quantizer threshold and the preferred communication direction concluded from investigating PCRLB are truly valid. In Fig. 4.7, numerical results of MSE at the FC with different correlation coefficients are displayed. One can easily see that the MSE, as same as the PCRLB in Fig. 4.3, is always minimized at $\gamma = 0$. Meanwhile, in Fig. 4.8, the simulations of MSE with different power ratios are visualized. By comparing it to Fig. 4.4, one can observe the same conclusion that the preferred communication direction is always from the node with the lower SNR to the higher one regardless of the correlation coefficient.

Fig. 4.9 illustrates the MMSE under different SNR at the worse node (σ_1^2), with fixed SNR at the fusion center (σ_2^2). One can see that as σ_1^2/σ_2^2 becomes larger, even small positive correlation can be exploited for better estimation performance. This is because the particular point $\rho = \frac{\sigma_2}{\sigma_1}$, at which the MSE reaches its maximum is closer to 0 when σ_1^2 increases.

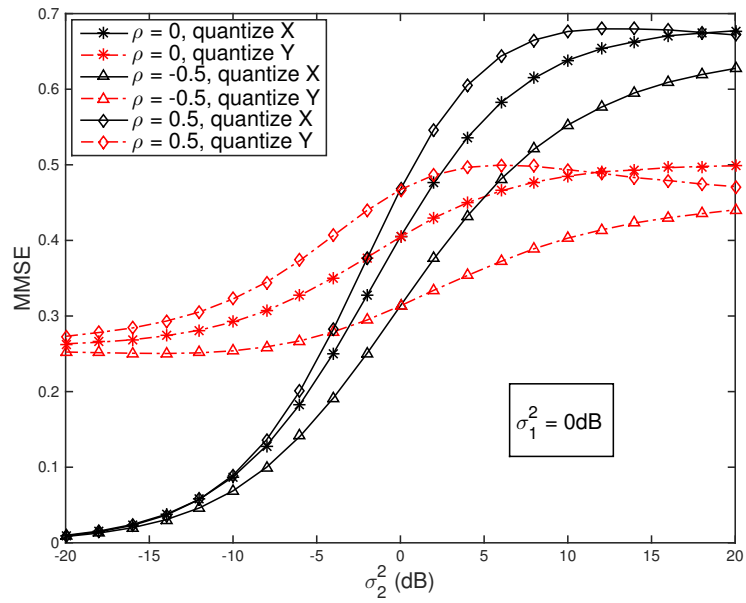


Fig. 4.8: Estimation MSE comparison with $\rho = -0.5, 0, 0.5$

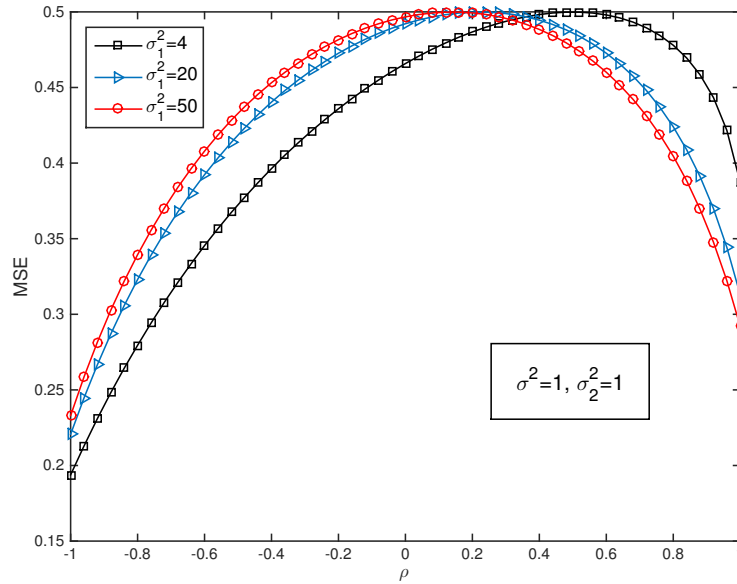


Fig. 4.9: Estimation MSE with different SNR at the worse node

4.4 Summary

In this chapter, we focused on decentralized estimation in a two-node tandem network with correlated Gaussian noises. The objective is to strive for a better understanding of the effect of data correlation on the estimation performance. With the Gaussian model, we first established the optimality of single threshold quantizer on local observations in maximizing the FI at the fusion center. This enables us to determine the optimal communication direction, which is from the node with lower SNR to the other one. Finally, different correlation regimes are characterized that have different ramifications with regard to their impacts on the estimation performance compared with that of independent observations.

A natural extension is to study a system involving more than two nodes. For the centralized case, the observation now becomes

$$\mathbf{y} = \theta + \mathbf{w}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma),$$

where Σ is the covariance matrix of the additive Gaussian noise. The joint FI to be

$$J(\mathbf{Y}; \theta) = -E \left[\frac{\partial^2}{\partial \theta^2} \log p(\mathbf{y}|\theta) \right] = \sum_{i,j} \tilde{\sigma}_{ij}, \quad (4.19)$$

where $\tilde{\sigma}_{i,j}$ stands for the (i, j) -th entry of Σ^{-1} . Clearly, for a general covariance matrix, the study of the impact of data correlation on the estimation performance is rather cumbersome. However, for special covariance matrix structures where correlation can be quantified using a single or a few parameters, systematic study is indeed possible. One such example is that all variables have identical pairwise correlation coefficient ρ . For this special case, similar results can be obtained that parallels the bivariate Gaussian noise case for a two-node system. The result is provided in Appendix D.

CHAPTER 5

ROOM SHAPE RECOVERY VIA A SINGLE MOBILE ACOUSTIC SENSOR WITHOUT TRAJECTORY INFORMATION

This chapter focus on the indoor room shape recovery using a single mobile acoustic sensor. In the problem settings, a polygonal room and a mobile node with co-located loudspeaker and microphone are employed, and only the first-order echoes are used to reconstruct a 2-D room geometry.

The room shape recovery problem has been extensively examined in the literature (see [14–19] and references therein). However, in those works, either a loudspeaker or microphone array is used, or high-order echoes are required; while in this section, the proposed algorithm uses only first-order echoes collected from a co-located loudspeaker/microphone sensor for room shape reconstruction. The reason that only first-order echoes are used is largely driven by practical considerations: higher-order echoes are often significantly weaker and of poor resolution, leading to unreliable time of arrival estimate. On the other hand, instead of assuming a single measurement point, we assume a mobile node (e.g., a mobile phone) that may take measurements at different locations. We show that in the ab-

sence of any knowledge of the measurement locations, first-order echoes are sufficient to recover a wide class of polygonal room shapes. We also investigate its limitation and identify the class of room shapes that are impossible to recover with the proposed system. This points to further research direction where additional information/measurements, such as the trajectory of the moving sensor, are required, which will be discussed in the next chapter.

5.1 Problem Formulation

5.1.1 Image Source Model

The basic technique employed in this chapter is the classic model in acoustics/optics, namely the image source model (ISM) [54–57]. This model converts a source inside a room into multiple ones outside the boundaries [55], each corresponding to an image of the original one. These are referred to as the image sources. The model is sketched in Fig.5.1, in which the polygon is encompassed by four edges W_1 - W_4 , and the source is located at O . Then S_1 - S_4 are called first-order images (e.g. the reflective path $O \rightarrow R_3 \rightarrow O$), while S_{12} and S_{21} are two examples of second-order images, indicating that they are the images over two distinct boundaries (the reflective path $O \rightarrow R_{12} \rightarrow R_{21} \rightarrow O$). The ISM helps us characterize the room impulse response (RIR), which we discuss below.

5.1.2 Mapping between ISM and RIR

In this section, we introduce the signal model and formulate the problem of room shape reconstruction in a 2-D space. Suppose the room to be measured is a polygon with N walls, denoted as W_j , $j = 1, \dots, N$. Let us consider a pair of co-located loudspeaker and microphone that moves on a set of points $\{O_i\}_{i=1}^S$, where each O_i is located at (x_i, y_i) . At each point, a probing signal $s(t)$ is transmitted from the loudspeaker, resulting in a series of echoes within the room. Define the ideal room impulse response at a particular point O_i

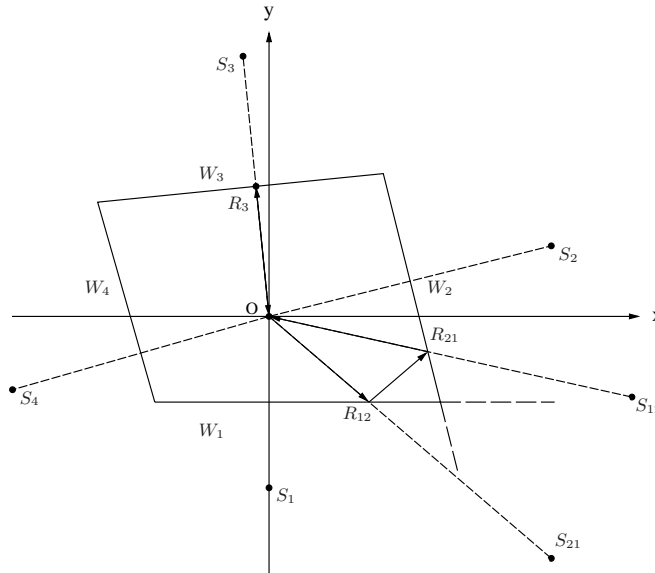


Fig. 5.1: Image Source Model

to be

$$h_i(t) = \sum_{n=0}^{\tilde{N}} a_{i,n} \cdot \delta(t - t_{i,n}), \quad (5.1)$$

where $a_{i,n}$ refers to the attenuation of the direct path ($n = 0$) or the reflective path between O_i and $S_{i,n}$, an image source of O_i , and the number of echoes \tilde{N} is generally greater than the number of walls N , as there may be second or even higher-order echoes involved. The parameters of interest in this model are the elapsed times $t_{i,n}$'s, i.e. the TOAs, and they can be obtained from the received signal $r_i(t)$, which is characterized by $r_i(t) = s(t) * h_i(t)$. Notice that for room shape recovery, the TOAs, not the exact form of $h_i(t)$ is needed. As such, broadband acoustic signals such as chirp signals are often candidate waveforms because of its good time domain resolution.

Assume that the speed of sound is known and invariant, one can map the TOAs with the room geometry. For the first-order echoes, denote the speed of sound as c and then the distance from O_i to W_j is

$$\frac{c \cdot t_{i,j}}{2}. \quad (5.2)$$

In [18], it was shown that the first-order TOAs measured at one single point is not enough to

uniquely recover a room, but combined first- and second-order measurements are sufficient. On the other hand, room shapes can be recovered solely based on the first-order echoes if multiple distributed microphones are used [15]. In this chapter, we demonstrate that by using a co-located and *mobile* loudspeaker and microphone, only first-order echoes are needed to reconstruct a class of 2-D room shape.

Echo labeling is an important issue that refers to the mapping between the RIR and room geometry, e.g., the k -th TOA in two different RIRs measured at two separate points may not correspond to the same wall. To simplify the problem, we assume that the measurement point moves within a substantially small range compared to the room size, such that the echoes arrive in the same order for all points.

5.2 Room Shape Reconstruction

In this section, we first establish that first-order TOAs are sufficient to retrieve a triangular room if three distinct sets of RIRs are provided. A simple algorithm to recover triangular room shapes is provided, which is subsequently extended to cope with more general polygonal shapes. Limitations of the proposed approach is also identified; we show that the first-order echo measurements alone has inherent limitation in dealing with parallelograms.

5.2.1 Identifiability by First-Order TOAs

We begin with a lemma that states that triangles can be recovered via three sets of first-order RIRs measured at three distinct locations without any information about the geometry of the three measurement points.

Lemma 5.1. *For any given triangle \mathcal{T} , suppose there are three interior points that are not on a straight line. Let \mathbf{R} be the distance matrix with R_{ij} being the distance between the i -th point and the j -th edge. Then \mathbf{R} uniquely defines the triangle with probability 1.*

Proof. The detailed proof is given in Appendix E. □

5.2.2 Algorithm for Triangle Reconstruction with Only First-order TOAs

The construction algorithm utilizes the distance matrix. For the sake of simplicity, let W_1 be on the x -axis, and one end of it be the origin. We denote the column vectors \mathbf{c}_i and \mathbf{r}_j as the i -th column and the j -th row of the transposed distance matrix \mathbf{R}^T . Under these settings, the two other walls are represented as two lines:

$$W_2 : y = a_2 \cdot x + b_2$$

$$W_3 : y = a_3 \cdot x.$$

The intercept of W_3 is 0 as it passes through the origin, and for each point, we have $y_i = R_{i1}$. Hence, we can map the first-order TOAs to the room geometry through six equations:

$$\frac{|a_2 x_i - R_{i1} + b_2|}{\sqrt{a_2^2 + 1}} = R_{i2}, \quad i = 1, 2, 3 \quad (5.3)$$

$$\frac{|a_3 x_i - R_{i1}|}{\sqrt{a_3^2 + 1}} = R_{i3}, \quad i = 1, 2, 3 \quad (5.4)$$

The solutions to (5.3) and (5.4) are:

$$a_3 = \pm \sqrt{\left(\frac{2BC}{D^2 - B^2 - C^2} \right) - 1}, \quad (5.5)$$

$$a_2 = \frac{Ba_3}{B + kC\sqrt{a_3^2 + 1}}, \quad (5.6)$$

$$b_2 = \frac{kA\sqrt{a_3^2 + 1}}{B + kC\sqrt{a_3^2 + 1}}, \quad (5.7)$$

$$x_i = \frac{R_{i1} + k\sqrt{a_3^2 + 1} \cdot R_{i3}}{a_3} \quad (5.8)$$

where $k = \text{sign}(a_3)$. Let $[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] = [\mathbf{c}_2, \mathbf{c}_3, \mathbf{c}_1]^T - \mathbf{R}^T$, then $A = \mathbf{r}_2^T(\mathbf{r}_1 \times \mathbf{r}_3)$, $B = \mathbf{r}_2^T \mathbf{u}_1$,

$$C = \mathbf{r}_2^T \mathbf{u}_3 \text{ and } D = \mathbf{r}_3^T \mathbf{u}_1.$$

From (5.5)-(5.8) we see that there are two sets of possible solutions; choosing the correct one depends on the angle formed by W_1 and W_3 , i.e., whether it is sharp or obtuse. It is not hard to discriminate between them, since the incorrect set of recovered points and edges are not able to reproduce the true distance matrix. For example, assume a_3 is actually negative, then in the set with $a_3 > 0$, the recovered distance between the reconstructed i -th point and W_2 is denoted as

$$D_{i2} = \frac{(A + BR_{i3} - CR_{i1}) \sqrt{a_3^2 + 1}}{\sqrt{(Ba_3)^2 + (B + C\sqrt{a_3^2 + 1})^2}}. \quad (5.9)$$

However, the true distance R_{i2} satisfies

$$R_{i2} = -\frac{(A + BR_{i3} - CR_{i1}) \sqrt{a_3^2 + 1}}{\sqrt{(Ba_3)^2 + (B - C\sqrt{a_3^2 + 1})^2}}, \quad (5.10)$$

which is not equal to D_{i2} . Hence, we have the following algorithm to recover a triangle.

Algorithm 5.1 Triangle Recovery

- 1: set $k = 1$
 - 2: calculate a_3, a_2, b_2 and x_i using (5.5)-(5.8)
 - 3: calculate $\mathbf{d}_2 = [D_{12}, D_{22}, D_{32}]^T$ using (5.9)
 - 4: **if** $\mathbf{d}_2 \neq \mathbf{r}_2$ **then**
 - 5: set $k = -1$ and go to step 2
 - 6: **end if**
-

5.2.3 Polygon Reconstruction by First-Order TOAs

We have shown that first-order RIRs are sufficient to recover triangular room shapes. Unfortunately, the result does not hold for arbitrary polygonal shapes. In particular, we demonstrate below that first-order RIRs alone are not sufficient to recover parallelograms.

Lemma 5.2. *For any given rectangle \mathcal{R} , there is an infinite number of parallelograms having the same distance matrix \mathbf{R} with \mathcal{R} , regardless of the number of measurement points.*

Proof. Suppose the edges of a rectangle \mathcal{R} are counter-clockwise ordered as $y = 0$, $x = L$, $y = H$ and $x = 0$. Then the first-order distance vectors must satisfy the following set of conditions:

$$\mathbf{r}_1 + \mathbf{r}_3 = H\mathbf{1}, \quad (5.11)$$

$$\mathbf{r}_2 + \mathbf{r}_4 = L\mathbf{1}, \quad (5.12)$$

where $\mathbf{1}$ is an all one column vector. It is easy to find a parallelogram enclosed by four lines: $y = 0$, $y = ax - \text{sign}(a)\sqrt{1+a^2}L$, $y = H$ and $y = ax$, such that its first-order TOAs also satisfy conditions (5.11) and (5.12). Since the slope a is an arbitrary real number, there are an infinite number of parallelograms that can not be distinguished from \mathcal{R} by first-order TOAs. \square

If the room shape is truly a rectangle, then Algorithm 1 can be easily adopted to recover the room shape. Suppose we randomly select three edges as W'_1 , W'_2 and W'_3 , then there must be a pair of parallel edges, therefore two cases can happen: (1) W'_1 is parallel to W'_2 , and (2) W'_2 is parallel to W'_3 . Note that we do not allow W'_1 parallel to W'_3 because the algorithm assumes that W'_1 intersects W'_3 at the origin (if $a_3 = 0$ is recovered, we can rearrange the vectors). Both cases are recovered by the algorithm as a shape with one perpendicular and two parallel edges, and the fourth one will later be recovered to enclose the shape as a rectangle.

The above reconstruction, however is still subject to the inherent ambiguity between the rectangle and any of the parallelograms that has the same \mathbf{R} matrix. The following corollary gives a positive result that shows that the first-order RIRs are sufficient to recover a wide class of 2-D room shapes.

Corollary 5.1. *For any polygon, if there is a subset of its edges forming a triangle, then it can be correctly and uniquely recovered solely based on its first-order RIRs.*

Proof. The proof is straightforward. Algorithm 5.1 fails only when the selected three edges can not form a triangle, i.e., two edges are parallel. Conversely, if there is a triangle, then we can correctly recover it, as well as the remaining edges. \square

5.2.4 Algorithm of Polygon Reconstruction with Higher-order TOAs

Suppose there are only first-order TOAs included in \mathbf{R} , we can reconstruct the polygon by repeatedly applying Algorithm 5.1 until all \mathbf{r}_j are examined. Specifically, we apply Algorithm 5.1 on the starting subset $\{\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3\}$ and keep the recovered a_3 and x_i 's. Then we apply Algorithm 5.1 on $\{\mathbf{r}_1, \mathbf{r}_k, \mathbf{r}_3\}$ for all $k = 4, \dots, N$, while we use the recorded a_3 and x_i 's to compute (a_k, b_k) . Finally, we need to check if the recovered distances $\mathbf{d}_k = [D_{1k}, D_{2k}, D_{3k}]^T$ match the true distances \mathbf{r}_k . If not, we need to rearrange the vectors of \mathbf{R} and start over again; this is necessary even for certain shapes that contain at least two parallel edges such as a regular hexagon. The algorithm will fail only if the room shape is a parallelogram. However, the \mathbf{R} matrix needs to contain only that of first-order echoes, thus the algorithm needs to be modified if higher-order echoes are included in the distance matrix.

Instead of \mathbf{R} , we use $\tilde{\mathbf{R}}$ to denote the distance matrix generated from the received echoes, which may contain higher-order TOAs, and its dimension is $3 \times \tilde{N}$. Let us first consider applying Algorithm 5.1 on $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$. There are three different cases.

- $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ does not result in a valid triangle

This happens when some subset of $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ does not correspond to first-order echos, then (5.5) may not have a real number solution. This is easy to fix: one can simply rearrange $\tilde{\mathbf{R}}$ until a triangle can be formed.

- $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ results in parallel edges or a valid but wrong triangle

If any two vectors in $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ correspond to parallel edges, such as in a regular hexagon, Algorithm 5.1 may mistakenly recover a right angle. Also, if any vector in the set does not correspond to a first-order echo, the recovered a_3 and x_i 's are wrong. In both cases, the subsequently recovered (a_k, b_k) are also wrong; nevertheless, if an $\tilde{\mathbf{r}}_k$ corresponds to an edge parallel to any one in $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$, the recovered distance still match the true distances, i.e. $\mathbf{d}_k = \tilde{\mathbf{r}}_k$.

- $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ results in a correct triangle

If all vectors in $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ come from first-order echoes without any parallel edges, then the $\tilde{\mathbf{r}}_k$'s correspond to first-order echoes will generate (a_k, b_k) such that $\mathbf{d}_k = \tilde{\mathbf{r}}_k$, but higher-order $\tilde{\mathbf{r}}_k$'s will not.

The discussion suggests that we need to pre-process the distance matrix $\tilde{\mathbf{R}}$ to pair all parallel edges and separate them into two sets, then recover (a_k, b_k) as mentioned above, and finally do a distance mismatch check to eliminate the higher-order TOAs and verify the reconstruction. Given the outline of algorithm, we can see that the polygons with more than three non-parallel edges and those with more than two pairs of parallel edges are able to be uniquely reconstructed. For the first case, if $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ is correctly selected (all come from first-order echoes), there is at least one \mathbf{r}_k for $k = 4, \dots, \tilde{N}$ which can pass the distance check to validate the starting subset; otherwise, no $\tilde{\mathbf{r}}_k$ can pass the check. Similarly, in the second case, more than two pairs of parallel edges guarantees a correct starting subset, and it follows the same reason as the first case. However, if a polygon has one or two pairs of parallel edges, it may not be correctly recovered due to inherent limitation of the problem settings. We have the following lemma to identify the polygons which can be uniquely recovered.

Lemma 5.3. *Suppose a polygon has N edges, including M pairs of parallel ones. The polygon can be correctly and uniquely recovered either when $N - M \geq 4$, or $M \geq 3$ if only first-order information are entirely available.*

Proof. The first condition generalizes the first case discussed above. Since we only take one of the two ‘parallel’ sets into account, then recovering the polygon is equivalent to recovering another one with $N - M$ non-parallel edges. As we have explained that a polygon with more than three non-parallel edges can be recovered, $N - M \geq 4$ is a sufficient condition for recoverability. The second condition is straightforward by seeing that at least three pairs of parallel edges guarantees a correct starting set when the algorithm is applied. \square

Given Lemma 5.3, the polygon that satisfies neither one of the two conditions has $M = 2, N = 4$ or 5 , or $M = 1, N = 4$, in another word, the trapezoid, parallelogram and parallelogram with one corner cut by a fifth edge can not be recovered by only using the first-order echoes. Consider that non-rectangular parallelogram-shaped room is very rare in practical, one can directly apply the algorithm to recover a rectangle. However, additional information, such as the geometry of the measurement points, is required to recover trapezoids, which will be discussed in the next chapter.

Another special case is that $M = 0, N = 3$, i.e., a triangle. We have shown in Sec. 5.2.2 that if only first-order TOAs are included in the distance matrix \mathbf{R} , then a triangle can be correctly recovered with probability 1. However, if high-order information is present, choosing a starting set $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ and then verifying other $\tilde{\mathbf{r}}_k$ by distance match check will lead the algorithm into an infinite loop. This can be seen from the fact that even the starting set is correctly chosen, no other $\tilde{\mathbf{r}}_k$ can pass the following distance match check, and the algorithm will automatically reject the current combination. The same happens to the case that the starting set is not correctly chosen.

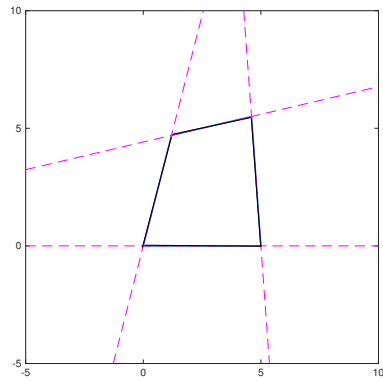
We have the algorithm given in Algorithm 6.1. Some numerical results are illustrated in Fig. 5.2. From the figures we can see that the reconstruction is subject to rotation and reflection ambiguity, and this is predictable as the coordinates of the measurement points are unknown.

Algorithm 5.2 Polygon Recovery

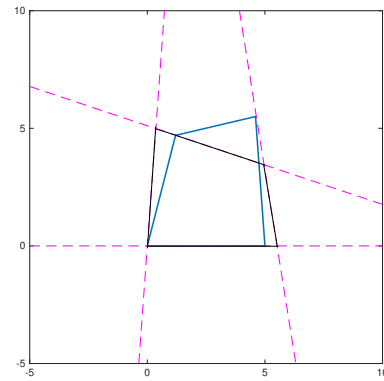
```

1: pair all ‘parallel’ vectors, put them into two separate sets  $V_1$  and  $V_2$ , and the rest in  $V_3$ 
2: rearrange  $\tilde{\mathbf{r}}$ : put  $V_1$  in the front, and  $V_2$  in the end
3: if  $M == 0$  or  $M \geq 3$  then
4:   apply Algorithm 5.1 on  $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_2, \tilde{\mathbf{r}}_3\}$ , record  $a_3$  and  $x_i$ 
5:   apply Algorithm 5.1 on  $\{\tilde{\mathbf{r}}_1, \tilde{\mathbf{r}}_k, \tilde{\mathbf{r}}_3\}$ , check if  $\mathbf{d}_k = \tilde{\mathbf{r}}_k$ 
6:   if  $\mathbf{d}_k \neq \tilde{\mathbf{r}}$  for all  $k$  then
7:     rearrange  $V_3$  and go to step 4
8:   else
9:     done.
10:  end if
11: else
12:  apply step 4-10 on  $V_1 \cup V_3$ ,
13:  if step 12 is done then
14:    recover the edges corresponding to  $V_2$ , done.
15:  else(repeats more than a maximum number of times)
16:    if  $M == 1$  then
17:      additional information is required.
18:    else( $M == 2$ )
19:      recover a rectangle, done.
20:    end if
21:  end if
22: end if

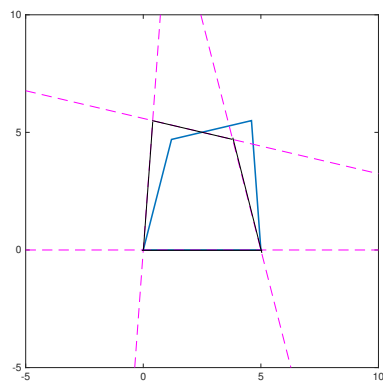
```



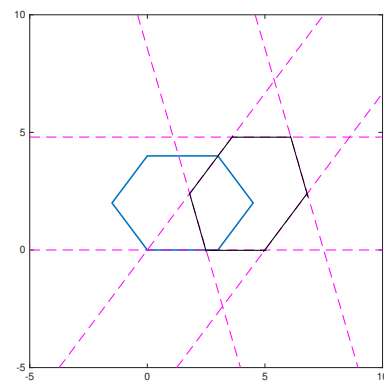
(a)



(b)



(c)



(d)

Fig. 5.2: original (blue) and recovered (dash purple/black) polygons: (a) an irregular quadrilateral, (b) a rotation, (c) a reflection, (d) a hexagon

5.3 Summary

While the proposed algorithm suffices to recover a wide class of a polygonal shapes through only first-order RIRs, additional information or measurement will be needed in order for the algorithm to work for all possible polygonal shapes. Restricting to only first-order RIRs, a natural direction is to add additional information with regard to the measurement points so that, together with the first-order RIRs, unique recovery (subject to reflection and rotation ambiguity) is guaranteed for all polygonal shapes. Specifically, we will show in the next chapter that with partial information about the trajectory of the moving sensor, all convex polygonal shapes can be recovered.

CHAPTER 6

ROOM SHAPE RECOVERY VIA A SINGLE MOBILE ACOUSTIC SENSOR WITH TRAJECTORY INFORMATION

In the previous chapter, we propose an algorithm to recover 2-D convex polygonal room shapes using a single mobile sensor with co-located loudspeaker and microphone. In the proposed algorithm, the mobile acoustic sensor need to move around the room and measure the TOAs at least at three distinct and non-collinear locations. Only the entire set of the first-order TOAs are required, while no information about the sensor trajectory, such as moving path length and the turning angle, is assumed to be known. It has shown in the last chapter that when the entire first-order and a subset of higher-order TOAs are included in the distance matrix used to recover the room shape, the proposed algorithm can not cope with triangle, trapezoid, parallelogram, and parallelogram with one corner cut by a fifth edge.

In the present chapter, we still consider the problem of reconstructing 2-D room shape with only the first-order echoes, but with information about the trajectory of the moving sensor, i.e., the geometry of the measurement points. Given the full geometry of the mea-

surement points, i.e., the distance between each pair, it is straightforward to establish that first order echoes are sufficient to recover any convex polygonal shapes. This however is a very strong assumption that often requires human intervention, e.g., actually measuring distances between measurement points. However, many mobile devices are capable of measuring its own path length when moving from one point to another. This chapter investigates the possibility of reconstructing 2- D room shape with path length of consecutive points measured by the mobile device. This weaker assumption, compared with the knowledge of complete geometry of measurement points, makes it feasible to achieve autonomous room shape recovery.

The present chapter establishes that the above approach is indeed feasible. That is, given the knowledge of path lengths between consecutive measurement points, one can recover arbitrary convex room shape by using only first-order echoes collected at three non-collinear points in the room. Algorithmic procedure that handles measurement noise and the presence of higher-order echoes is also proposed and experimental results are presented to validate the approach.

6.1 System Model

The basic model used in the present chapter is the same as the last one, say, the ISM. We repeat some important notations in the last chapter as follows. We use \mathbf{R} to denote the distance matrix, where R_{ij} represents the distance between the i -th measurement point and the j -th wall, and \mathbf{c}_i and \mathbf{r}_j are the i -th column and the j -th row of \mathbf{R}^T , respectively. In the experiment, we denote the number of echoes received at the i -th measurement point as N_i .

6.1.1 Geometry

Consider a convex planar K -polygon. As shown in Fig. 6.1, a mobile device with co-located microphone and loudspeaker emits pulses and receives echoes at $\{O_i\}_{i=1}^3$. Without

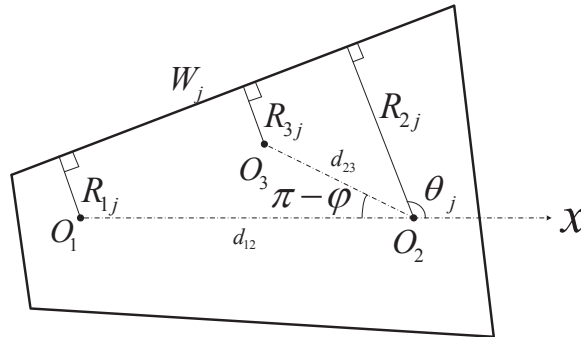


Fig. 6.1: A mobile device is employed to measure the geometry of a room. The mobile device collects echoes at O_1 , O_2 and O_3 successively. The distances between the consecutive measurement points are d_{12} and d_{23}

loss of generality, we assume that O_1 is the origin, O_2 lies on the x -axis, and O_3 lies above the x -axis. Let $\varphi = (\pi - \angle O_1 O_2 O_3) \in (0, \pi)$ and the length of $O_1 O_2$ and $O_2 O_3$ be denoted by d_{12} and d_{23} , respectively.¹ The angle between the x -axis and the normal vector of the j -th wall is denoted by θ_j .

From Fig. 6.1, it is straightforward to show that

$$(R_{2j} - R_{1j}) + d_{12} \cos \theta_j = 0, \quad (6.1)$$

and

$$d_{23} \cos(\theta_j - \varphi) + (R_{3j} - R_{2j}) = 0. \quad (6.2)$$

If however the R_{ij} 's contain higher-order echoes, with probability 1, (6.1) and (6.2) do not hold simultaneously.

Clearly given d_{12} and d_{23} and if φ can also be estimated from the measurement data, then the geometry of the measurement points can be recovered. In this case, first-order echoes are sufficient to recover room geometry.

¹If $\pi \in (0, 2\pi)$, i.e. we do not have control of where to place O_3 , then the reconstruction is subject to reflection ambiguity (c.f. Theorem 6.1).

6.2 Recovery with Known Distances and Unknown Path Direction

If for all measurement points, the one-to-one mapping $f_i : \{R_{ij}\}_{j=1}^K \mapsto \mathbf{c}_i$ is known, then $\{\theta_j\}_{j=1}^K$ can be obtained by (6.1) and the room shape is determined. However since echoes may arrive in different orders at different O_i 's and \mathbf{c}_i may contain higher-order echoes if $N_i > K$, f_i is unknown. Then θ_j 's are also unknown. Therefore we need a way to both rule out higher-order echoes and find the correct combination of the first-order echoes. We can then estimate θ_j 's and the room shape.

Define $\alpha_{jj'} = -\frac{R_{2j}-R_{1j'}}{d_{12}}$ and $\beta_{jj'} = -\frac{R_{3j'}-R_{2j}}{d_{23}}$. For simplicity we denote α_{jj} and β_{jj} by α_j and β_j , respectively. From (6.1) and (6.2), we have

$$\theta_j = \pm \arccos \alpha_j \quad \text{and} \quad \theta_j - \varphi = \pm \arccos \beta_j, \quad (6.3)$$

Thus, there are four possible sign combinations for a given j ,

$$\theta_j = \arccos \alpha_j \quad \text{and} \quad \theta_j - \varphi = \arccos \beta_j \quad (6.4)$$

$$\theta_j = \arccos \alpha_j \quad \text{and} \quad \theta_j - \varphi = -\arccos \beta_j \quad (6.5)$$

$$\theta_j = -\arccos \alpha_j \quad \text{and} \quad \theta_j - \varphi = \arccos \beta_j \quad (6.6)$$

$$\theta_j = -\arccos \alpha_j \quad \text{and} \quad \theta_j - \varphi = -\arccos \beta_j. \quad (6.7)$$

We first give the definition of feasibility as follows.

Definition 1. *Given a room \mathcal{R} and a location O , we say O is feasible if the co-located device at O can receive all the first-order echoes of a signal emitted at O .*

Then we have the following lemmas stating the identifiability of convex polygons using first-order echoes. We discuss the cases of grouped (echoes for every measurement points are sorted in a same order with respect to the walls) and ungrouped echoes, separately.

Lemma 6.1. *Suppose O_1, O_2 and O_3 are feasible and not collinear. Given the correct echo combination, with probability 1, there exist exactly two sign combinations such that (6.1) and (6.2) hold simultaneously for all j 's if φ and the direction of both $\overrightarrow{O_1O_2}$ and $\overrightarrow{O_2O_3}$ are randomly chosen. The two possible sign combinations have opposite signs for φ and all θ_j 's and correspond to reflection of each other.*

Proof. Assume that the ground truth of the polygon is (6.4) for all $j \in \{1, \dots, K\}$. Note that (6.4) implies that (6.7) holds for $\theta'_j = -\theta_j$ and $\varphi' = -\varphi < 0$ for all j , which is the reflection of the room.

Suppose multiple sign combinations hold for a wall. Without loss of generality, let $j = 1$. From (6.4) we have

$$\varphi = \arccos \alpha_1 - \arccos \beta_1. \quad (6.8)$$

Assume that one of the following equations also holds,

$$\varphi = -\arccos \alpha_1 - \arccos \beta_1 \quad (6.9)$$

$$\varphi = \arccos \alpha_1 + \arccos \beta_1 \quad (6.10)$$

$$\varphi = -\arccos \alpha_1 + \arccos \beta_1. \quad (6.11)$$

Then we have the following three cases:

1. If (6.8) and (6.9) hold, we must have $\theta_1 = 0$ which implies that O_1O_2 is perpendicular to the first wall, and $\varphi = -\arccos \beta_1$.
2. If (6.8) and (6.10) hold, we must have $\arccos \beta_1 = 0$, which implies that O_2O_3 is perpendicular to the first wall.
3. If (6.8) and (6.11) hold, we must have $\varphi = 0$, which contradict with the assumption that O_1, O_2 and O_3 are not collinear.

With probability 1, the first two cases do not occur since both φ and directions of $\overrightarrow{O_1O_2}$ and $\overrightarrow{O_2O_3}$ are randomly chosen.

If a subset of (6.5)-(6.7) holds for j and j' simultaneously, then we must have $(\theta_j, \theta_{j'}) \in \{\theta_j = 0, \theta_j = \varphi, \varphi = 0\} \times \{\theta_{j'} = 0, \theta_{j'} = \varphi, \varphi = 0\}$, which again, do not occur due to randomly chosen measurement points. Similarly, it can be shown that for more than two walls, (6.4) would imply none of (6.5)-(6.7) holds for all walls. \square

Lemma 6.2. *Given incorrect echo combinations, with probability 1, there exists no solution to (6.1) and (6.2).*

Proof. We illustrate the proof by considering only the case of $K = 4$. The result can be easily extended to $K = 3$ and $K > 4$.

We assume that the ground truth is (6.4) for all j . First consider parallelograms. The distances between O_i ($i = 1, 2, 3$) and the four walls satisfy

$$R_{11} + R_{12} = R_{21} + R_{22} = R_{31} + R_{32} = a, \quad (6.12)$$

and

$$R_{13} + R_{14} = R_{23} + R_{24} = R_{33} + R_{34} = b. \quad (6.13)$$

We can see that for certain echo combinations, pairs of $\{\alpha_{jj'}, \beta_{jj'}\}$, $j, j' \in \{1, 2, 3, 4\}$ are dependent. Consider for example the echo combination resulting in $\{\alpha_{12}, \alpha_{21}, \alpha_{34}, \alpha_{43}\}$ and $\{\beta_{12}, \beta_{21}, \beta_{34}, \beta_{43}\}$. Since $\alpha_{12} + \alpha_{21} = 0$, $\alpha_{34} + \alpha_{43} = 0$, $\beta_{12} + \beta_{21} = 0$ and $\beta_{34} + \beta_{43} = 0$, we have

$$\arccos(\alpha_{21}) = \pi \pm \arccos(\alpha_{12})$$

$$\arccos(\alpha_{43}) = \pi \pm \arccos(\alpha_{34})$$

$$\arccos(\beta_{21}) = \pi \pm \arccos(\beta_{12})$$

$$\arccos(\beta_{43}) = \pi \pm \arccos(\beta_{34}).$$

Thus (6.3) reduces to two equations

$$\varphi = \pm \arccos(\alpha_{12}) \pm \arccos(\beta_{12})$$

$$\varphi = \pm \arccos(\alpha_{34}) \pm \arccos(\beta_{34}).$$

With probability 1, these two equations do not hold simultaneously as α_{12}, β_{12} are independent of α_{34}, β_{34} due to randomly chosen measurement points. Other possible cases of incorrect combination always have at least two equations with independent choice of α and β . Hence no solution can be found for those instances.

Suppose a combination of echoes is chosen such that we have $\alpha_{jj'}$ and $\beta_{jj''}$ ($i \neq i', i \neq i''$). For rooms with no more than one pair of parallel walls, almost surely no echo combination other than the correct one can make (6.4) holds for all j . This is because for those rooms, at least one of (6.12) and (6.13) does not hold. Thus some $\alpha_{jj'}$'s and $\beta_{jj''}$'s are not dependent since $R_{1j'}, R_{2j}$ and $R_{3j''}$ are randomly chosen from $\mathbf{c}_1, \mathbf{c}_2$ and \mathbf{c}_3 , respectively.

Therefore only the correct combination of first-order echoes can satisfy (6.4) for all walls. □

Note that the distances between the device and both the ceiling and the floor can be ruled out by Lemma 6.2.

Given Lemma 6.1 and Lemma 6.2, we have the following result on the identifiability of any convex polygonal room by using only first-order echoes.

Theorem 6.1. *One can recover, with probability 1, any convex planar K -polygon subject to reflection ambiguity, by using the first-order echoes received at three random points in the feasible region, with known d_{12} and d_{23} and unknown $\varphi \in (0, 2\pi)$.*

Remark: The room shape is subject to reflection ambiguity for $\varphi \in (0, 2\pi)$. If, however, we can limit $\varphi \in (0, \pi)$, the recovered shape is unique, i.e. not subject to ambiguity.

In the presence of noise, however, \mathbf{c}_i is subject to measurement errors. Hence φ solved

from (6.3) for different j 's are not identical. A straightforward practical algorithm that handles the measurement errors is given below:

Algorithm 6.1 All Convex Polygon Recovery

- 1: set $\tilde{N} = \min \{N_1, N_2, N_3\}$
 - 2: **for** $K = 3 : \tilde{N}$ **do**
 - 3: choose K entries from each of \mathbf{c}_i , $i = 1, 2, 3$
 - 4: **for** $k = 1 : \dots, \binom{N}{K}^3 (K!)^2$ **do**
 - 5: **for** $j = 1 : K$ **do**
 - 6: compute $\varphi_j^k = \pm \arccos \alpha_j \pm \arccos \beta_j$
 - 7: **end for**
 - 8: **end for**
 - 9: choose and record the echo combination with the smallest variance of φ_j^k for the given K
 - 10: **end for**
 - 11: choose the recorded echo combination with the largest K with the variance of φ_j^k less than a threshold
 - 12: estimate θ_j 's using the obtained combination of echoes and reconstruct the polygon
-

The following lemma states that the knowledge of both d_{12} and d_{23} is necessary for reconstructing any convex polygons by first-order echoes.

Lemma 6.3. *If either d_{12} or d_{23} is missing, then 1) parallelogram can not be reconstructed. 2) Non-parallelogram can be reconstructed subject to reflection ambiguity.*

The proof of the part 1) is to construct a counterexample while 2) can be established in a manner similar to that of Lemma 6.2.

6.3 Experimental result

In this section, we discuss our experiments conducted in a shoebox classroom.

6.3.1 Experiment Setup

We use a laptop as a microphone and a HTC M8 phone as our loudspeaker. As the loudspeaker of the cell phone is not omnidirectional and power limited, we place the speaker

of the cell phone towards each wall to ensure the corresponding first-order echo is strong enough. Note that the loudspeaker will record first-order as well as some higher-order ones.

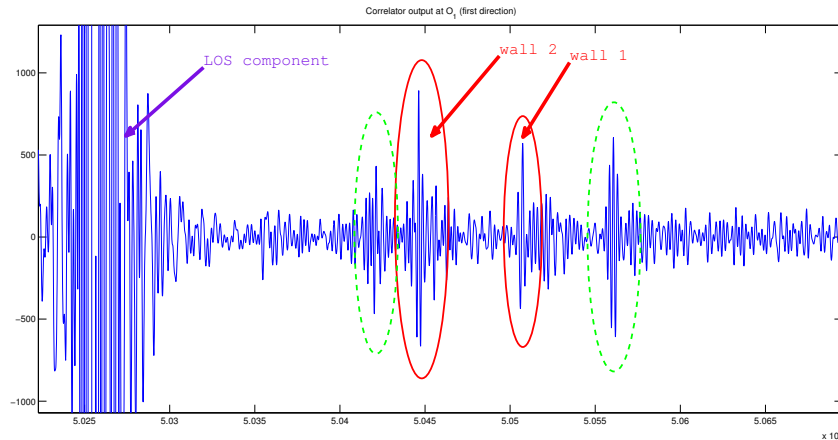


Fig. 6.2: Correlator output at O_1 towards the first wall. Peaks with solid ellipses correspond to true walls. Peaks with dash ellipses correspond to either noise or higher-order echoes

6.3.2 Signal Type

A chirp signal linearly sweeping from 30Hz to 8kHz is emitted by the cell phone. The sample rate at the receiver is 96kHz. It has been shown in [58, 59] that if the input chirp signal is correlated with its windowed version, the output may resemble a delta function. Our simulation and experiment results show that the candidate distances are obtained by correlating the received signals with its triangularly windowed version outperforms the correlator output using the original one. Fig. 6.2 is a sample path of the correlator output collected in the room where this experiment is conducted.

6.3.3 Room Shape Reconstruction Experiment

The proposed approach is verified by experiment in which d_{12} and d_{23} are measured by tape measure. Even if some elements of c_i have measurement errors up to 10cm, the room

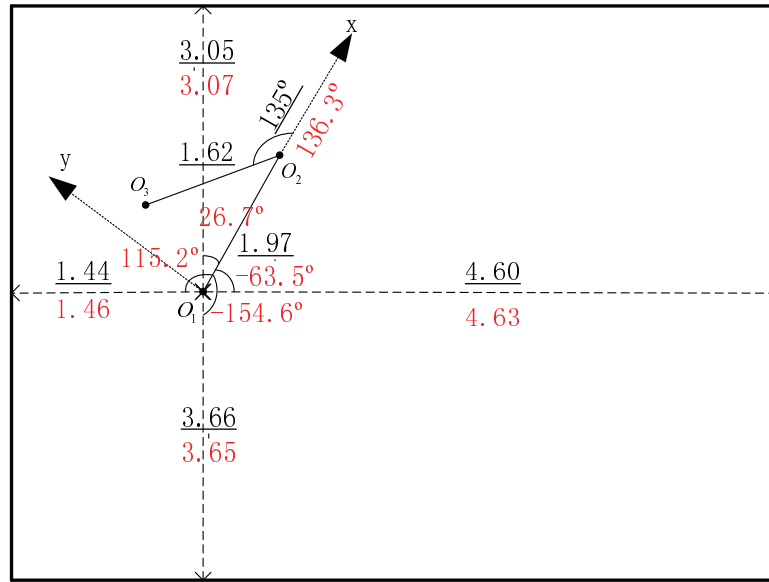


Fig. 6.3: Comparison between the ground truth (black underlined) and experiment result (red)

can be recovered with negligible error by the proposed algorithm if the subset of c_i corresponding to first-order echoes are known. In the presence of higher-order echoes, the proposed algorithm performs poorly when the variance criterion is the only criterion used to determine the correct combination of echoes. Since most rooms are regular, we add a heuristic constraint: all the angles of two adjacent walls are between 30° and 150° . An interesting phenomenon is that sometimes the proposed algorithm is unable to provide the correct room shape, but φ is always estimated close to the true value. Therefore, one can use the algorithm in Sec. 6.2 to obtain φ and then reconstruct the room shape independently with full knowledge of the geometry information of the measurement points. The comparison between the reconstruction result and the ground truth is illustrated in Fig. 6.3.

6.4 Summary

In this chapter, we make progress in room shape reconstruction using only first-order echoes. Specifically, we established that given partial information about the sensor trajec-

tory, i.e., the distances between consecutive measurement points, any 2- D convex polygon can be reconstructed. In the presence of noise, a simple algorithm is devised that is effective in recovering the room shape even if the higher-order echoes are present.

However, it can be seen that the time complexity of the proposed algorithm is shown to be $\mathcal{O}\left(\sum_{k=3}^N k \binom{N}{k}^3 (k!)^2\right)$, implying that it is an NP hard problem. Therefore, there are further research can be done to rule out a large part of the echo combinations to reduce the complexity of the problem.

CHAPTER 7

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

7.1 Concluding Remarks

In nowadays communication and electronic systems, multi-functional sensors are widely used. They are able to sense and collect information from the surrounding environment, perform relatively simple computation locally, and also collaborate with each other utilizing their communication capability. In this thesis, we addressed problems in two separate types of applications of such sensors. In the first part, we investigate the decentralized inference with dependent observations in wireless sensor networks. Two different problems are studied in this part, one is about the cooperative spectrum sensing system using multiple sensing nodes, and the other one is regarding a decentralized estimation system with correlated Gaussian noises. In the second part, a single mobile sensor bundled with co-located microphone and loudspeaker is used to recover 2-D convex polygonal room shape based on the acoustic response of the room.

Chapter 3 is dedicated to study the energy detection in spectrum sensing system. In this chapter, the optimality (or sub-optimality) of energy detection in two types of spec-

trum sensing is examined. For the single node system, the energy detector is proved to be optimal for most signaling, while for the only case of not being optimal, it is shown to be substantially close to the theoretically optimal one. For cooperative spectrum sensing case, a recently proposed framework for distributed inference with dependent observations (reviewed in Chapter 2) is employed to help establish the optimality of the energy detector for some signaling, while in other cases, the problem remains difficult to solve.

Chapter 4 is focused on decentralized estimation with correlated Gaussian noises in a tandem network. The impact of noise correlation on decentralized estimation performance is investigated. For a tandem network, the optimality of SFQ in the sense of maximizing FI is established; furthermore, with correlated Gaussian noises, the SFQ reduces to threshold quantizer on local observations, which is subsequently optimal. Additionally, it is shown to be better off quantizing the worse node and having the FC at the better node for all correlation regimes. Finally, different correlation regimes in terms of their impact on the decentralized estimation performance are identified. The regimes include the well known cases of negatively and highly positively correlated noise benefiting estimation due to noise cancellation, and another positive correlation regime which induces redundancy.

Chapter 5 considers a practical problem in room shape recovery using only first-order echoes. In this problem, a single mobile sensor, equipped with co-located loudspeaker and microphone, as well as a possible internal motion sensors, is deployed and TOAs are measured at multiple locations to recover the room shape. In this chapter, no motion sensor is assumed, thus no trajectory information about the mobile sensor is available. The uniqueness of the mapping between the first-order TOAs and the room shape is identified when at least three distinct non-collinear measurement locations are used. It is also shown that in the absence of higher-order TOAs, all convex polygonal shapes except for parallelogram can be uniquely and correctly recovered the proposed algorithm, while more information is required to tackle the parallelogram case.

Chapter 6 addresses the same problem with that in *Chapter 5*, however, partial trajectory information of the mobile sensor is accessible as the motion sensors are able to measure the path length between consecutive measurement locations. In this case, the mapping between the first-order TOAs and any convex polygonal room shape is proved, and an algorithm is proposed to reconstruct room shape in the presence of higher-order echoes and noise.

7.2 Future Research Directions

We conclude this thesis by listing a few future research directions.

- For energy detection in cooperative spectrum sensing, it is possible that energy detector is optimum for more general cases, e.g., the detection of Gaussian signal in Gaussian channels with more than one sample. This, however, requires generalization of the framework developed in [20] to more general hidden variables. Additionally, other more realistic channel fading models can be considered, including non-Rayleigh fading channels, as well as slow fading channels (i.e., correlation of fading channels in time).
- For the room shape reconstruction with partial trajectory information of the mobile sensor, the proposed algorithm is able to deal with any convex polygonal shape, with the expense of time complexity as high as $\mathcal{O}\left(\sum_{k=3}^N k \binom{N}{k}^3 (k!)^2\right)$. This is a high time complexity that sensors, or even personal computers, typically can not process locally in a tolerable time. Therefore, it is certainly worth improving the algorithm time complexity by excluding as many as possible the candidate shapes in advance.

APPENDIX A

DISTRIBUTION OF PSK AND QAM

SIGNALS OVER RAYLEIGH FADING

CHANNEL

In the Rayleigh fading channel, the complex fading coefficient h can be written in a polar form as $h = |h| \cdot e^{j\theta_h}$. It is complex Gaussian distributed with the variance σ^2 , i.e.

$$p(h) = \frac{1}{\pi\sigma^2} e^{-\frac{|h|^2}{\sigma^2}},$$

which means that the real part $\mathcal{R}\{h\} \triangleq h_r$ and the imaginary part $\mathcal{I}\{h\} \triangleq h_i$ are both Gaussian distributed with zero-mean and variance $\frac{\sigma^2}{2}$, and they are independent.

For a particular PSK symbol s_m , the output signal through the Rayleigh fading channel is $h \cdot s_m$, then it can be further rewritten as $|h| \cdot e^{j(\theta_h + \theta_m)}$, where θ_m is a deterministic quantity. Then it is not difficult to verify that the real and imaginary part of $h \cdot s_m$ are also independent and Gaussian distributed. Furthermore, since $|h \cdot s_m| = |h|$, the faded signal $h \cdot s_m$ has exactly the same distribution as h . Therefore, the output signal $x(n) = h(n)s(n)$

has the distribution

$$p(x) = \sum_m \pi_m p(h \cdot s_m) = \frac{1}{\pi \sigma^2} e^{-\frac{|h|^2}{\sigma^2}},$$

which is also exactly identical with that of the fading h .

For a particular QAM symbol, similar to the PSK case, the distribution of $h \cdot s_m = |h| r_m \cdot e^{j(\theta_h + \theta_m)}$ is complex Gaussian with zero-mean and scaled variance $r_m^2 \sigma^2$. What makes it different from PSK is that r_m are not identical for all constellation points, so the distribution of the output signal $x(n) = h(n) * s(n)$ is

$$p(x) = \sum_m \pi_m p(h \cdot s_m) = \sum_m \frac{1}{\pi r_m^2 \sigma^2} e^{-\frac{|h|^2}{r_m^2 \sigma^2}},$$

which is a mixed Gaussian.

APPENDIX B

PROOF OF LEMMA 4.3

Our goal is to prove that for any $\gamma \in \mathbf{R}$,

$$J(\psi(X, \gamma); \theta|Y) + J(Y; \theta) > J(X; \theta) + J(\psi(Y, \gamma); \theta|X). \quad (\text{B.1})$$

With the Gaussian model for θ and the additive noises, (B.1) can be expanded into (B.2).

$$\begin{aligned}
& \frac{1}{\sigma_2^2} + \frac{(1 - \frac{\sigma_1}{\sigma_2}\rho)^2}{2\pi(1 - \rho^2)\sigma_1^2} \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \frac{e^{-\frac{(y-\theta)^2}{2\sigma_2^2}}}{\sqrt{2\pi\sigma_2^2}} \frac{e^{-\left(\frac{\gamma-\theta-\frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right)^2}}{Q\left(\frac{\gamma-\theta-\frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right) \left(1 - Q\left(\frac{\gamma-\theta-\frac{\sigma_1}{\sigma_2}\rho(y-\theta)}{\sqrt{1-\rho^2}\sigma_1}\right)\right)} dy d\theta \\
& > \frac{1}{\sigma_1^2} + \frac{(1 - \frac{\sigma_2}{\sigma_1}\rho)^2}{2\pi(1 - \rho^2)\sigma_2^2} \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \frac{e^{-\frac{(x-\theta)^2}{2\sigma_1^2}}}{\sqrt{2\pi\sigma_1^2}} \frac{e^{-\left(\frac{\gamma-\theta-\frac{\sigma_2}{\sigma_1}\rho(x-\theta)}{\sqrt{1-\rho^2}\sigma_2}\right)^2}}{Q\left(\frac{\gamma-\theta-\frac{\sigma_2}{\sigma_1}\rho(x-\theta)}{\sqrt{1-\rho^2}\sigma_2}\right) \left(1 - Q\left(\frac{\gamma-\theta-\frac{\sigma_2}{\sigma_1}\rho(x-\theta)}{\sqrt{1-\rho^2}\sigma_2}\right)\right)} dx d\theta.
\end{aligned} \quad (\text{B.2})$$

Proof. Rewrite the two conditional FIs as:

$$\begin{aligned}
J_1 &\triangleq J(\psi(X, \gamma); \theta|Y) \\
&\stackrel{t=\frac{y-\theta}{\sigma_2}}{=} C_1 \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \cdot \eta\left(\frac{\gamma - \theta - \sigma_1 \rho t}{\sqrt{1 - \rho^2 \sigma_1}}\right) dt d\theta, \\
J_2 &\triangleq J(\psi(Y, \gamma); \theta|X) \\
&\stackrel{t=\frac{x-\theta}{\sigma_1}}{=} C_2 \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \cdot \eta\left(\frac{\gamma - \theta - \sigma_2 \rho t}{\sqrt{1 - \rho^2 \sigma_2}}\right) dt d\theta,
\end{aligned}$$

where $C_1 = \frac{(1 - \frac{\sigma_1 \rho}{\sigma_2})^2}{2\pi(1 - \rho^2)\sigma_1^2}$, $C_2 = \frac{(1 - \frac{\sigma_2 \rho}{\sigma_1})^2}{2\pi(1 - \rho^2)\sigma_2^2}$, $f(\theta)$ is the pdf of the parameter θ , and $\phi(\cdot)$ is the pdf of standard normal distribution. Thus it suffices to prove $J_2 - J_1 < \frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}$.

Using the fact that $\max \eta(t) = 4$, one can further bound the difference of the two conditional FIs as in (B.3).

$$\begin{aligned}
J_2 - J_1 &= (C_2 - C_1) \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \cdot \eta\left(\frac{\gamma - \theta - \sigma_2 \rho t}{\sqrt{1 - \rho^2 \sigma_2}}\right) dt d\theta \\
&\quad + C_1 \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \cdot \left(\eta\left(\frac{\gamma - \theta - \sigma_2 \rho t}{\sqrt{1 - \rho^2 \sigma_2}}\right) - \eta\left(\frac{\gamma - \theta - \sigma_1 \rho t}{\sqrt{1 - \rho^2 \sigma_1}}\right) \right) dt d\theta \\
&= \frac{1}{2\pi} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right) \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \cdot \eta\left(\frac{\gamma - \theta - \sigma_2 \rho t}{\sqrt{1 - \rho^2 \sigma_2}}\right) dt d\theta \\
&\quad + C_1 \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \cdot \left(\eta\left(\frac{\gamma - \theta - \sigma_2 \rho t}{\sqrt{1 - \rho^2 \sigma_2}}\right) - \eta\left(\frac{\gamma - \theta - \sigma_1 \rho t}{\sqrt{1 - \rho^2 \sigma_1}}\right) \right) dt d\theta \\
&< \frac{2}{\pi} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right) + C_1 \int_{-\infty}^{\infty} f(\theta) \int_{-\infty}^{\infty} \phi(t) \left(\eta\left(\frac{\theta - \gamma + \sigma_2 \rho t}{\sqrt{1 - \rho^2 \sigma_2}}\right) - \eta\left(\frac{\theta - \gamma + \sigma_1 \rho t}{\sqrt{1 - \rho^2 \sigma_1}}\right) \right) dt d\theta.
\end{aligned} \tag{B.3}$$

The task is now equivalent to showing that the second term in (B.3) is non-positive. Let

$a_1 \triangleq \frac{\theta - \gamma}{\sqrt{1 - \rho^2 \sigma_1}}$, $a_2 \triangleq \frac{\theta - \gamma}{\sqrt{1 - \rho^2 \sigma_2}}$, $s \triangleq \frac{\rho t}{\sqrt{1 - \rho^2}}$. The inner integral in (B.3) is

$$\frac{\sqrt{1 - \rho^2}}{|\rho|} \cdot \int_{-\infty}^{\infty} \frac{e^{-\frac{(1 - \rho^2)s^2}{2\rho^2}}}{\sqrt{2\pi}} \cdot (\eta(a_2 + s) - \eta(a_1 + s)) ds \tag{B.4}$$

Define

$$q(s) = \eta(a_2 + s) - \eta(a_1 + s).$$

We can show that $q(s)$ is an odd function at $s = -\frac{a_1+a_2}{2}$, i.e. $q(s - \frac{a_1+a_2}{2}) = -q(-s - \frac{a_1+a_2}{2})$. Without loss of generality, consider $\theta > \gamma$, which implies $a_2 > a_1 > 0$. Consider the following different regions of s in terms of the effect on the sign of $q(s)$.

- $s < -a_2$, then $a_1 + s < a_2 + s < 0$. In this region, $\eta(t)$ is a monotonically increasing function, thus

$$q(s) = \eta(a_2 + s) - \eta(a_1 + s) > 0.$$

- $-a_2 < s < -\frac{a_1+a_2}{2}$. This condition implies that $a_1 + s < a_1 - \frac{a_1+a_2}{2} < 0$ and $0 < a_2 + s < a_2 - \frac{a_1+a_2}{2}$, thus

$$\begin{aligned} \eta(a_2 + s) &\stackrel{(1)}{>} \eta\left(a_2 - \frac{a_1 + a_2}{2}\right) \\ &= \eta\left(a_1 - \frac{a_1 + a_2}{2}\right) \stackrel{(2)}{>} \eta(a_1 + s). \end{aligned}$$

The inequality (1) holds as $\eta(t)$ decreases monotonically when $t > 0$ and similarly for (2). The equality follows as $\eta(t)$ is symmetric at $t = 0$. Therefore, in this region,

$$q(s) = \eta(a_2 + s) - \eta(a_1 + s) > 0.$$

Thus $q(s) > 0$ when $s < -\frac{a_1+a_2}{2}$. Due to the odd symmetry of $q(s)$ about $-\frac{a_1+a_2}{2}$, we can easily obtain that $q(s) < 0$ when $s > -\frac{a_1+a_2}{2}$.

If $\theta < \gamma$, then $a_2 < a_1 < 0$, the discussions are similar and the reverse inequality holds.

The results can be summarized as follows

$$\begin{aligned} q(s) \cdot \left(s + \frac{a_1 + a_2}{2}\right) &< 0, \text{ for } \theta > \gamma, \\ q(s) \cdot \left(s + \frac{a_1 + a_2}{2}\right) &> 0, \text{ for } \theta < \gamma. \end{aligned}$$

$$\begin{aligned}
& \frac{\sqrt{1-\rho^2}}{|\rho|} \left(\int_{-\infty}^{-\frac{a_1+a_2}{2}} \frac{e^{-\frac{(1-\rho^2)s^2}{2\rho^2}}}{\sqrt{2\pi}} q(s) + \int_{-\frac{a_1+a_2}{2}}^{+\infty} \frac{e^{-\frac{(1-\rho^2)s^2}{2\rho^2}}}{\sqrt{2\pi}} q(s) \right) ds \\
& \stackrel{(1)}{=} \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}|\rho|} \left(\int_{-\infty}^{-\frac{a_1+a_2}{2}} e^{-\frac{(1-\rho^2)s^2}{2\rho^2}} q(s) ds + \int_{-\infty}^{-\frac{a_1+a_2}{2}} e^{-\frac{(1-\rho^2)(-(a_1+a_2)-r)^2}{2\rho^2}} q(-(a_1+a_2)-r) dr \right) \\
& \stackrel{s=r}{=} \frac{\sqrt{1-\rho^2}}{\sqrt{2\pi}|\rho|} \left(\int_{-\infty}^{-\frac{a_1+a_2}{2}} \left(e^{-\frac{(1-\rho^2)s^2}{2\rho^2}} - e^{-\frac{(1-\rho^2)(-(a_1+a_2)-s)^2}{2\rho^2}} \right) q(s) ds \right). \tag{B.5}
\end{aligned}$$

Therefore, we can split the integral in (B.4) into two parts and rearrange the terms in (B.5).

In (B.5), the first equation follows by changing the variable $r = -(a_1 + a_2) - s$. We now analyze the integral in (B.5) for two cases:

- $\theta - \gamma > 0$.

As $s < -\frac{a_1+a_2}{2}$ in the integral in (B.5), $q(s)$ is always positive as discussed before.

Meanwhile, it is not difficult to see that $s^2 > (-(a_1 + a_2) - s)^2$, thus $e^{-\frac{(1-\rho^2)s^2}{2\rho^2}} - e^{-\frac{(1-\rho^2)(-(a_1+a_2)-s)^2}{2\rho^2}} < 0$. Hence, the integral is negative.

- $\theta - \gamma < 0$.

Under this condition, $q(s)$ is always negative as discussed. Moreover, as $a_2 <$

$a_1 < 0$ when $\theta - \gamma < 0$, we have $s^2 < (-(a_1 + a_2) - s)^2$, and thus $e^{-\frac{(1-\rho^2)s^2}{2\rho^2}} - e^{-\frac{(1-\rho^2)(-(a_1+a_2)-s)^2}{2\rho^2}} > 0$. Hence the integral is also negative.

Therefore, the integral is always less than 0. Plug this result back into (B.3), the difference of the two conditional FI's is bounded by:

$$J_2 - J_1 < \frac{2}{\pi} \left(\frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2} \right) < \frac{1}{\sigma_2^2} - \frac{1}{\sigma_1^2}.$$

This proves (B.1) hence

$$J(\psi(X, \gamma), Y; \theta) > J(X, \psi(Y, \gamma); \theta).$$

Although we assume that the parameter θ is Gaussian distributed, the entire proof is independent of the distribution of θ . Hence, for arbitrary θ , within the framework of single threshold quantizer, it is always better to impose it at the sensor with the worse noise, though the single threshold quantizer is not always optimal for arbitrary parameter. \square

APPENDIX C

PROOF OF LEMMA 4.2

To prove lemma 4.2, we only need to show that

$$\frac{1}{p(t)} = e^{t^2} Q(t)(1 - Q(t))$$

is a monotonic increasing function over $(0, \infty)$, i.e.,

$$\begin{aligned} & \frac{d}{dt} e^{t^2} Q(t)(1 - Q(t)) \\ & = e^{t^2} \left(2tQ(t)(1 - Q(t)) - \frac{1}{\sqrt{2\pi}}(1 - 2Q(t))e^{-\frac{t^2}{2}} \right) > 0. \end{aligned} \quad (\text{C.1})$$

Since it is continuous at $t = 1$, we can prove it separately on two sets, $(1, \infty)$ and $[0, 1]$.

- Monotonicity of $p(t)$ over $(1, \infty)$

We only need to show that $h(t) = e^{t^2} Q(t)$ is monotonically increasing function on $(1, \infty)$, i.e., $h'(t) > 0$, as $(1 - Q(t))$ increases monotonically in this region. To prove it, we need to use an inequality

$$Q(t) > \frac{t}{\sqrt{2\pi}(1 + t^2)} \cdot e^{-\frac{t^2}{2}}. \quad (\text{C.2})$$

Taking derivative of $h(t)$, we have

$$\begin{aligned} h'(t) &= 2te^{t^2}Q(t) - e^{t^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \\ &\stackrel{(1)}{>} \frac{e^{\frac{t^2}{2}}}{\sqrt{2\pi}} \left(\frac{2t^2}{1+t^2} - 1 \right) > 0, \forall t > 1. \end{aligned}$$

The inequality (1) follows Eq. (C.2). Hence, $p(t)$ decreases over $(1, \infty)$.

- Monotonicity of $p(t)$ over $[0, 1]$

Observe Eq. (C.1), it is a quadratic form of $Q(t)$, though the coefficients are dependent on t . We can construct a function

$$f(t) = -2t \cdot x^2 + \left(2t + \frac{2}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \right) \cdot x - \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}.$$

Then the roots of $f(t) = 0$ are

$$\begin{aligned} x_1(t) &= \frac{1}{2} - \frac{\sqrt{2\pi} t e^{\frac{t^2}{2}}}{2(\sqrt{2\pi} t^2 e^{t^2} + 1 + 1)}, \\ x_2(t) &= \frac{1}{2} + \frac{\sqrt{2\pi} t e^{\frac{t^2}{2}}}{2(\sqrt{2\pi} t^2 e^{t^2} + 1 - 1)}. \end{aligned}$$

If we can prove that $x_1(t) \leq Q(t) \leq x_2(t)$ over $(0, 1)$, then the inequality in Eq. (C.1) is satisfied. First, since $0 \leq Q(t) \leq \frac{1}{2}$ over $(0, 1)$, then $Q(t) \leq x_2(t)$ always holds. To show $x_1(t) \leq Q(t)$, we first check two boundary points: $x_1(0) = 0.5 = Q(0)$, and $x_1(1) = 0.1066 < 0.1587 = Q(1)$. The derivative of $r(t) \triangleq Q(t) - x_1(t)$ is given as

$$r'(t) = -\frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} + \frac{\sqrt{2\pi}}{2} \frac{(1+t^2)e^{\frac{t^2}{2}}}{\sqrt{2\pi} t^2 e^{t^2} + 1(\sqrt{2\pi} t^2 e^{t^2} + 1 + 1)},$$

so we can check the derivative on the two boundary points:

$$\begin{aligned} r'(t)|_{t=0} &= -\frac{1}{\sqrt{2\pi}} + \frac{\sqrt{2\pi}}{2} > 0, \\ r'(t)|_{t=1} &= -\frac{1}{\sqrt{2\pi e}(\sqrt{2\pi e} + 1)} < 0, \end{aligned}$$

then there must be at least one t_0 such that $r'(t_0) = 0$. Furthermore, if there is only one such t_0 , combined with $r(0) = 0$, $r(1) > 0$, it is easy to show that $r(t) > 0$ over $(0, 1)$. Identifying t_0 is equivalent to look for the root of the following equation:

$$\pi e^{t^2} \cdot (t^2 - 1)^2 = 2. \quad (\text{C.3})$$

Let $x \triangleq t^2$, and define $g(x) = \pi e^x \cdot (x - 1)^2$. Then we have

$$\begin{aligned} g(0) &= \pi > 2, \\ g(1) &= 0 < 2, \\ g'(x) &= \pi e^x \cdot (x^2 - 1) < 0, \quad x \in (0, 1), \end{aligned}$$

which implies that there exists one and only one $x_0 \in (0, 1)$ such that $g(x_0) = 2$, and consequently one and only one $t_0 \in (0, 1)$ such that $r'(t_0) = 0$, which means $r(t) = Q(t) - x_1(t) > 0, t \in [0, 1]$.

Overall, the function $p(t)$ is monotonically decreasing over $(0, \infty)$, and increasing over $(-\infty, 0)$, where $t^* = 0$ gives its maximum $p^* = 4$. Also, $\lim_{t \rightarrow \pm\infty} p(t) = 0$, so-called the bell-shaped function.

APPENDIX D

CENTRALIZED ESTIMATION FOR A MULTI-NODE SYSTEM IN GAUSSIAN NOISES

Extension to the multi-sensor scenario is feasible under certain symmetric assumption of the correlation matrix which is included below. While this does not cover all possible covariance matrices of observations, the intention is to provide insight on how various correlation regimes might impact estimation performance.

Consider the case of a multi-node system with multivariate additive Gaussian noises that have identical pairwise correlation coefficient. The sensor observations can thus be written as

$$\mathbf{y} = \theta + \mathbf{w}, \quad \mathbf{w} \sim \mathcal{N}(\mathbf{0}, \Sigma),$$

where Σ is the covariance matrix taking on the following special structure

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \cdots & \rho\sigma_1\sigma_n \\ \vdots & \ddots & \vdots \\ \rho\sigma_n\sigma_1 & \cdots & \sigma_n^2 \end{bmatrix}.$$

It is a straightforward exercise to show that the joint Fisher information is

$$J(\mathbf{Y}; \theta) = -E \left[\frac{\partial^2}{\partial \theta^2} \log p(\mathbf{y}|\theta) \right] = \sum_{i,j} \tilde{\sigma}_{ij},$$

where $\tilde{\sigma}_{i,j}$ stands for the (i, j) -th entry of Σ^{-1} , the inverse of the covariance matrix Σ :

$$\Sigma^{-1} = \begin{bmatrix} \frac{1+(n-2)\rho}{(1-\rho)(1+(n-1)\rho)\sigma_1^2} & \cdots & \frac{-\rho}{(1-\rho)(1+(n-1)\rho)\sigma_1\sigma_n} \\ \vdots & \ddots & \vdots \\ \frac{-\rho}{(1-\rho)(1+(n-1)\rho)\sigma_n\sigma_1} & \cdots & \frac{1+(n-2)\rho}{(1-\rho)(1+(n-1)\rho)\sigma_n^2} \end{bmatrix}.$$

Therefore the joint FI can be expressed as

$$J(\mathbf{Y}; \theta) = \frac{1+(n-2)\rho}{(1-\rho)(1+(n-1)\rho)} \cdot \sum_{i=1}^n \frac{1}{\sigma_i^2} - \frac{\rho}{(1-\rho)(1+(n-1)\rho)} \cdot \sum_{i=1}^n \sum_{j \neq i} \frac{1}{\sigma_i \sigma_j}.$$

The contrasting case is when the noises are mutually independent, and the corresponding Fisher information is

$$J(U; \theta|Y) \Big|_{\rho=0} = \sum_{i=1}^n \frac{1}{\sigma_i^2}.$$

This serves as the baseline to study the impact of noise correlation on the estimation performance. Thus, the problem becomes comparing the following quantity to 0

$$d \triangleq J(\mathbf{Y}; \theta) - J(\mathbf{Y}; \theta) \Big|_{\rho=0} = \frac{(n-1)\rho^2 \sum_{i=1}^n \frac{1}{\sigma_i^2} - \rho \sum_{i=1}^n \sum_{j \neq i} \frac{1}{\sigma_i \sigma_j}}{(1-\rho)(1+(n-1)\rho)},$$

and then one can find the different regimes of ρ that has distinct impact on the estimation performance.

- $\rho = -\frac{1}{n-1}$. This value leads to a rank deficient covariance matrix, which means that in this case, the noises are linearly dependent. Therefore, perfect estimation can be achieved by complete noise cancellation, i.e., by a linear combiner with properly chosen coefficients.

- $-\frac{1}{n-1} < \rho < 0$. This is the regime of ρ such that data correlation may benefit estimation performance through partial noise cancellation. It is comparable to the negative correlation regime in bivariate Gaussian scenario.
- $0 < \rho < \frac{\sum_{i=1}^n \sum_{j \neq i} \frac{1}{\sigma_i \sigma_j}}{(n-1) \sum_{i=1}^n \frac{1}{\sigma_i^2}}$. In this regime d is negative, which means noise correlation leads to redundancy. Additionally, $\frac{\sum_{i=1}^n \sum_{j \neq i} \frac{1}{\sigma_i \sigma_j}}{(n-1) \sum_{i=1}^n \frac{1}{\sigma_i^2}}$ is the boundary point defined in a similar manner to that of the bivariate Gaussian case.
- $\frac{\sum_{i=1}^n \sum_{j \neq i} \frac{1}{\sigma_i \sigma_j}}{(n-1) \sum_{i=1}^n \frac{1}{\sigma_i^2}} \leq \rho < 1$. Positive high correlation is beneficial to estimation as it does for the bivariate case. In this case, noise cancellation more than compensating for signal power reduction.
- $\rho = 1$. Similar to the discussions in bivariate Gaussian noise scenario, perfect estimation can be achieved at this value if noise variance are not identical to each other.

Thus the study of the bivariate case (i.e., a two-node system) generalizes to the multi-variate case. Similar observations can be made for the decentralized system yet the analysis is rather cumbersome and one has to largely resort to numerical evaluation.

APPENDIX E

PROOF OF LEMMA 5.1

Proof. As depicted in Fig. E.1(a), the sides of T are counter-clockwisely labeled. Without any loss of generality, we fix O_1 to be the origin, and the first side of the triangle to be $y = -r_{11}$. To check if there is any triangle other than T such that there exist three points O'_i satisfying $\mathbf{r}'_i = \mathbf{r}_i$, we need to consider two cases: one side is tilted and two sides are tilted. We will consider them separately.

1. One side is tilted

Suppose W_3 stays intact and W_2 is being tilted, then O_2 has to keep at its original position, i.e. $O'_2 = O_2$. As shown in Fig. E.1(a), observe that W_2 is an external common tangent of the two circles centered at O_1 and O_2 with radii r_{12} and r_{22} , it is easy to show that the other external common tangent W'_2 is the only candidate satisfying $\mathbf{r}'_i = \mathbf{r}_i, i = 1, 2$. Therefore, two measure points are not enough to uniquely define a triangle. However, to keep $\mathbf{r}'_3 = \mathbf{r}_3$, O'_3 has to be on the line O_1O_2 . Hence, for three noncollinear points inside T , there is no triangle remaining \mathbf{r}_i while having only one side different from T .

2. Two sides are tilted

In this case, both W_2 and W_3 are being tilted. Now randomly fix W'_3 to be any

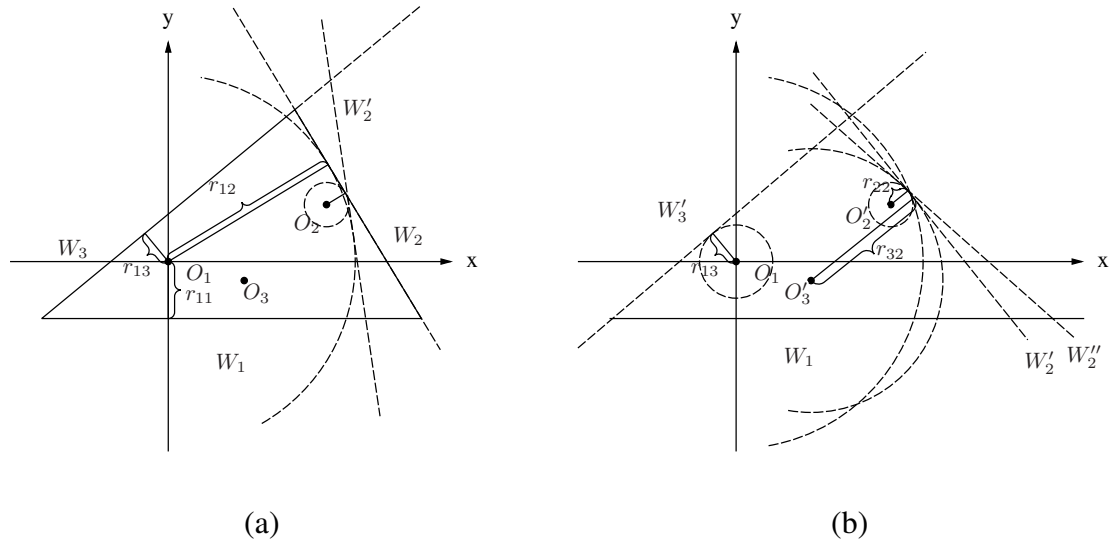


Fig. E.1: (a) one side is tilted, (b) two sides are tilted

line that is r_{13} apart from O_1 , and without any loss of generality, we choose it to be positive sloped and above the origin. Then W'_3 is represented as:

$$W'_3 : y = a_3 \cdot x + r_{13} \cdot \sqrt{a_3^2 + 1}.$$

Therefore, the other two points O'_2 and O'_3 can be easily determined as:

$$O'_2 = \left(\frac{(r_{21} - r_{11}) + (r_{23} - r_{13}) \cdot \sqrt{a_3^2 + 1}}{a_3}, r_{21} - r_{11} \right),$$

$$O'_3 = \left(\frac{(r_{31} - r_{11}) + (r_{33} - r_{13}) \cdot \sqrt{a_3^2 + 1}}{a_3}, r_{31} - r_{11} \right).$$

Now consider the second side W'_2 , which can be generally expressed by

$$W'_2 : a_2 \cdot x + b_2 \cdot y + c_2 = 0,$$

and the coefficients must satisfy the following equations to keep \mathbf{r}_2 unchanged:

$$|c_2| = r_{12} \quad (\text{E.1})$$

$$\left| \frac{a_2}{a_3} \left((r_{21} - r_{11}) + (r_{23} - r_{13}) \sqrt{a_3^2 + 1} \right) + b_2(r_{21} - r_{11}) + c_2 \right| = r_{22} \quad (\text{E.2})$$

$$\left| \frac{a_2}{a_3} \left((r_{31} - r_{11}) + (r_{33} - r_{13}) \sqrt{a_3^2 + 1} \right) + b_2(r_{31} - r_{11}) + c_2 \right| = r_{32} \quad (\text{E.3})$$

$$\sqrt{a_2^2 + b_2^2} = 1 \quad (\text{E.4})$$

Eq. (E.1) indicates that c_2 is a constant independent of a_2 and b_2 . As equivalent to linear equations of (a_2, b_2) , Eq. (E.2) - (E.3) have finite numbers of solutions, referring to the common tangents of the two circles centered at O'_2 and O'_3 with radii r_{22} and r_{32} . Meanwhile, Eq. (E.4) implies that (a_2, b_2) must be a point on the unit circle on the 2-dimensional plane of (a_2, b_2) . Since a_3 is randomly selected, then the intersection of two random lines falls on the unit circle has probability 0, because the boundary is of measure 0 in a 2-dimensional coordinate system. On the other hand, if a_3 is correctly selected, then there must be a single solution to Eq. (E.1) - (E.4), which is T . This case is illustrated in Fig. E.1(b), where W'_2 and W''_2 are two common tangents.

In conclusion, any triangle can be exactly recovered by measuring at three noncollinear points inside it, when first-order information is provided. \square

REFERENCES

- [1] R. Radner, “Team decision problems,” *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 857–881, 1962.
- [2] Pramod K Varshney, *Distributed detection and data fusion*, Springer Science & Business Media, 2012.
- [3] R. Viswanathan and P.K. Varshney, “Distributed detection with multiple sensors: part I - fundamentals,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, 1997.
- [4] J.N. Tsitsiklis, “Decentralized detection,” in *Advances in Statistical Signal Processing*. 1993, vol. 2, pp. 297–344, JAI Press, Greenwich, CT.
- [5] J.A. Gubner, “Distributed estimation and quantization,” *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1456–1459, 1993.
- [6] D. Blatt and A Hero, “Distributed maximum likelihood estimation for sensor networks,” in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Montreal, QC, Canada, May 2004, vol. 3, pp. iii–929–32 vol.3.
- [7] Federal Communications Commission, “Spectrum policy task force,” Nov. 2002.
- [8] Q. Zhao and B.M. Sadler, “A survey of dynamic spectrum access,” *IEEE Signal Processing Magazine*, vol. 24, no. 3, pp. 79–89, May. 2007.
- [9] D. Cabric, S.M. Mishra, and R.W. Brodersen., “Implementation issues in spectrum sensing for cognitive radios,” in *Proc. Asilomar Conference on Signals, Systems and Computers*, Nov. 2004, vol. 1, pp. 772 – 776.

- [10] A. Ghasemi and E.S. Sousa, “Collaborative spectrum sensing for opportunistic access in fading environments,” in *New Frontiers in Dynamic Spectrum Access Networks*, Nov. 2005, pp. 131–136.
- [11] Zhi-Quan Luo, “Universal decentralized estimation in a bandwidth constrained sensor network,” *IEEE Trans. Inf. Theory*, vol. 51, no. 6, pp. 2210–2219, 2005.
- [12] A. Ribeiro and G.B. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case,” *IEEE Trans. Signal Process.*, vol. 54, no. 3, pp. 1131–1143, 2006.
- [13] A. Ribeiro and G.B. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-part II: unknown probability density function,” *IEEE Trans. Signal Process.*, vol. 54, no. 7, pp. 2784–2796, 2006.
- [14] D. Ba, F. Ribeiro, Cha Zhang, and D. Florencio, “L1 regularized room modeling with compact microphone arrays,” in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, March 2010, pp. 157–160.
- [15] I. Dokmanic, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vertterli, “Acoustic echoes reveal room shape,” in *Proceedings of the National Academy of Sciences of the United States of America*, July 2013, vol. 110.
- [16] S. Tervo and T. Tossavainen, “3d room geometry estimation from measured impulse responses,” in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Kyoto, Japan, March 2012, pp. 513–516.
- [17] F. Antonacci, A. Sarti, and S. Tubaro, “Geometric reconstruction of the environment from its response to multiple acoustic emissions,” in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Dallas, TX, USA, March 2010, pp. 2822–2825.

- [18] I. Dokmanic, Y.M. Lu, and M. Vetterli, “Can one hear the shape of a room: The 2-d polygonal case,” in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Prague, Czech Republic, May 2011, pp. 321–324.
- [19] M. D. Plumbley, “Hearing the shape of a room,” in *Proceedings of the National Academy of Sciences of the United States of America*, 2013.
- [20] Hao Chen, Biao Chen, and P.K. Varshney, “A new framework for distributed detection with conditionally dependent observations,” *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1409–1419, March 2012.
- [21] R.S. Blum, S.A. Kassam, and H.V. Poor, “Distributed detection with multiple sensors: Part II - advanced topics,” *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, 1997.
- [22] Z. Chair and P. K. Varshney, “Optimal data fusion in multiple sensor detection systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-22, no. 1, pp. 98–101, Jan 1986.
- [23] E. Drakopoulos and C. C. Lee, “Optimum multisensor fusion of correlated local decisions,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 27, no. 4, pp. 593–606, Jul 1991.
- [24] M. Kam, Q. Zhu, and W. S. Gray, “Optimal data fusion of correlated local decisions in multiple sensor detection systems,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 28, no. 3, pp. 916–920, Jul 1992.
- [25] I. Y. Hoballah and P. K. Varshney, “Distributed bayesian signal detection,” *IEEE Trans. Inf. Theory*, vol. 35, no. 5, pp. 995–1000, Sep 1989.
- [26] J.N. Tsitsiklis and M. Athans, “On the complexity of decentralized decision making and detection problems,” in *IEEE Trans. Autom. Control*, May. 1985, vol. 30, pp. 440–446.

- [27] P.-N. Chen and A. Papamarcou, "Likelihood ratio partitions for distributed signal detection in correlated gaussian noise," *Proc. IEEE Intl. Symp. Inf. Theory*, p. 118, Oct. 1995.
- [28] P. Willett, P.F. Swaszek, and R.S. Blum, "The good, bad and ugly: distributed detection of a known signal in dependent gaussian noise," *IEEE Trans. Signal Process.*, vol. 48, no. 12, pp. 3266 – 3279, Dec. 2000.
- [29] H. Urkowitz, "Energy detection of unknown deterministic signals," *Proceedings of the IEEE*, vol. 55, no. 4, pp. 523 – 531, Apr. 1967.
- [30] F.F. Digham, M.S. Alouini, and M.K. Simon, "On the energy detection of unknown signals over fading channels," *IEEE Trans. Commun.*, vol. 55, no. 1, pp. 21 –24, Jan. 2007.
- [31] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *IEEE Communications Surveys Tutorials*, vol. 11, no. 1, pp. 116 –130, 2009.
- [32] A. Nuttall, "Some integrals involving the q-m function (corresp.)," *IEEE Trans. Inf. Theory*, vol. 21, no. 1, pp. 95–96, Jan 1975.
- [33] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*, vol. 2, Prentice Hall PTR, 1993.
- [34] D.J. Warren and P.K. Willett, "Optimum quantization for detector fusion: Some proofs, examples and pathology," *Journal of Franklin Institute*, vol. 336, pp. 323–359, Mar. 1999.
- [35] H. Chen, B. Chen, and P.K. Varshney, "A new framework for distributed detection with conditionally dependent observations," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1409 –1419, Mar. 2012.

- [36] R. Viswanathan, “A note on distributed estimation and sufficiency,” *IEEE Trans. Inf. Theory*, vol. 39, no. 5, pp. 1765–1767, Sep. 1993.
- [37] Harry L. Van Trees and Kristine L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*, Wiley-IEEE Press, 2007.
- [38] P. Tichavsky, C. H. Muravchik, and A. Nehorai, “Posterior Cramer-Rao bounds for discrete-time nonlinear filtering,” *IEEE Trans. Signal Process.*, vol. 46, no. 5, pp. 1386–1396, May 1998.
- [39] W.-M Lam and A.R. Reibman, “Design of quantizers for decentralized estimation systems,” *IEEE Trans. Commun.*, vol. 41, no. 11, pp. 1602–1605, 1993.
- [40] P. Venkitasubramaniam, G. Mergen, L. Tong, and A. Swami, “Quantization for distributed estimation in large scale sensor networks,” in *Proc. Intl. Conf. on Intelligent Sensing and Info. Process.*, Bangalore, India, 2005, pp. 121–127.
- [41] P. Venkitasubramaniam, L. Tong, and A. Swami, “Score-function quantization for distributed estimation,” in *Proc. Annual Conference on Information Sciences and Systems (CISS)*, Princeton, NJ, USA, 2006, pp. 369–374.
- [42] A. S. Behbahani, A. M. Eltawil, and H. Jafarkhani, “Decentralized estimation under correlated noise,” *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5603–5614, Nov 2014.
- [43] S. Liu, S. P. Chepuri, M. Fardad, E. Masazade, G. Leus, and P. K. Varshney, “Sensor selection for estimation with correlated measurement noise,” *IEEE Trans. on Signal Process.*, vol. 64, no. 13, pp. 3509–3522, July 2016.
- [44] John N. Tsitsiklis, “Decentralized detection,” in *In Advances in Statistical Signal Processing*. 1993, pp. 297–344, JAI Press.

- [45] J.N. Tsitsiklis, “Extremal properties of likelihood-ratio quantizers,” in *Proc. IEEE Conf. on Decision and Control*, Honolulu, HI, USA, 1990, pp. 2680–2685 vol.5.
- [46] Z.-B. Tang, K.R. Pattipati, and D.L. Kleinman, “A distributed M-ary hypothesis testing problem with correlated observations,” in *Proc. IEEE Conf. on Decision and Control*, Tampa, FL, USA, Dec 1989, pp. 562–568 vol.1.
- [47] Qing Yan and R.S. Blum, “Distributed signal detection under the Neyman-Pearson criterion,” *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1368–1377, May 2001.
- [48] J.D. Papastavrou and M. Athans, “On optimal distributed decision architectures in a hypothesis testing environment,” *IEEE Trans. Autom. Control*, vol. 37, no. 8, pp. 1154–1169, 1992.
- [49] E. Song, X. Shen, J. Zhou, Y. Zhu, and Z. You, “Performance analysis of communication direction for two-sensor tandem binary decision system,” *IEEE Trans. Inf. Theory*, vol. 55, no. 10, pp. 4777–4785, 2009.
- [50] E. Akofor and B. Chen, “On optimal fusion architecture for a two-sensor tandem distributed detection system,” in *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Austin, TX, USA, Dec 2013, pp. 129–132.
- [51] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*, vol. 1, Prentice Hall PTR, 1993.
- [52] R. Zamir, “A proof of the Fisher information inequality via a data processing argument,” *IEEE Trans. Inf. Theory*, vol. 44, no. 3, pp. 1246–1250, 1998.
- [53] G. Xu and B. Chen, “The sufficiency principle for decentralized data reduction,” in *Proc. IEEE Intl. Symposium on Info. Theory*, Boston, MA, USA, 2012, pp. 319–323.

- [54] J. B. Allen and D. A. Berkley, “Image method for efficiently simulating small-room acoustics,” in *Journal of the Acoustical Society of America*, April 1979, vol. 65(4), pp. 943–950.
- [55] J. Borish, “Extension of the image model to arbitrary polyhedra,” in *Journal of the Acoustical Society of America*, June 1984, vol. 75(6), pp. 1827–1836.
- [56] P. Peterson, “Simulating the response of multiple microphones to a single acoustic source in a reverberant room,” in *Journal of the Acoustical Society of America*, Nov. 1986, vol. 80(5), pp. 1527–1529.
- [57] E. Lehmann and A. Johansson, “Prediction of energy decay in room impulse responses simulated with an image-source model,” in *Journal of the Acoustical Society of America*, Jul. 2008, vol. 124(1), pp. 269–277.
- [58] Angelo Farina, “Simultaneous measurement of impulse response and distortion with a swept-sine technique,” in *Audio Engineering Society Convention 108*. Audio Engineering Society, 2000.
- [59] Guy-Bart Stan, Jean-Jacques Embrechts, and Dominique Archambeau, “Comparison of different impulse response measurement techniques,” *Journal of the Audio Engineering Society*, vol. 50, no. 4, pp. 249–262, 2002.

VITA

NAME OF AUTHOR: Fangrong Peng

PLACE OF BIRTH: Wuhan, China

DATE OF BIRTH: Nov. 17, 1986

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

Syracuse University, Syracuse, NY, United States

Chalmers University of Technology, Gothenburg, Sweden

Huazhong University of Science and Technology, Wuhan, China

DEGREES AWARDED:

M. S., 2014, Syracuse University, Syracuse, NY, United States

M. S., 2010, Chalmers University of Technology, Gothenburg, Sweden

B. S., 2008, Huazhong University of Science and Technology, Wuhan, China