

**ON THE USE OF LEAST FAVORABLE DISTRIBUTIONS TO FACILITATE THE DESIGN OF  
RANDOMLY DEPLOYED SENSOR DETECTION SYSTEMS**

by

Benedito José Barreto Fonseca Jr.

A dissertation submitted in partial fulfillment of  
the requirements for the degree of

Doctor of Philosophy

(Electrical Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2012

Date of final oral examination: 05/08/12

The dissertation is approved by the following members of the Final Oral Committee:

John A. Gubner, Professor, Electrical and Computer Engineering

Nigel Boston, Professor, Mathematics and Electrical and Computer Engineering

James Bucklew, Professor, Electrical and Computer Engineering

Robert Nowak, McFarland-Bascom Professor in Engineering, Electrical and Computer Engineering

Akbar Sayeed, Professor, Electrical and Computer Engineering



To my son, Gustavo.

## ACKNOWLEDGMENTS

One of the advantages of going back to the university after having worked in the profession for some years is that you go back with a clear picture of the knowledge and skills that you want to acquire. In my case, I decided to start a Ph.D. to improve my mathematical modeling skills and deepen my knowledge of probability theory and stochastic processes. When I mentioned this to a colleague of mine at Motorola, Raúl E. Sequeira, he told me that I ought to work with Prof. John A. Gubner. Raúl was right.

I thank Prof. John A. Gubner for teaching me invaluable lessons in probability theory and in how to rigorously build and analyze mathematical models. I am thankful for his patience and for his understanding of my peculiar situation, which involved the time limitations imposed by my concurrent work at Motorola, my family obligations, and the constraints imposed by living far from the university. I also thank Prof. Gubner for giving me the freedom to choose and guide my research topic while illuminating it with his advice and experience. I am also grateful for the many conversations in which he provided me with insightful professional advice. It was an honor and a privilege to work with him.

I also thank Prof. Timo Seppäläinen of the Mathematics department. As Jerry P. King rightfully describes in his book *“The Art of Mathematics,”* I, like many of us, entered graduate school thinking of mathematics as a tool for solving problems. Then, when I enrolled into my first course in analysis, as King’s words perfectly describe, I *“saw, for the first time, a professor who treated mathematics with reverence, who wrote symbols on the blackboard with great care as if they mattered as much as the information they contained. We heard a mathematical result described as elegant. And we saw that it was. It was a moment of great discovery.”* I was fortunate for being Prof. Seppäläinen’s student in my first course in analysis and in many of his other classes. His crystal-clear classes opened my eyes to the beauty of mathematics. I thank him for this lifelong gift.

I also thank Motorola Inc., now Motorola Mobility, for its continuous support of my Ph.D. studies, from the first to the last day. I also thank the many Motorolans who supported me, directly or indirectly, during this period. It is said that a great company is formed by great people. Motorola has shown itself to be a great company. I have had many colleagues and supervisors during my many Ph.D. years. All of them valued my pursuit of a higher education, fully supported me, and were understanding of the need for having to work part-time. I recall a conversation with one of my previous supervisors during the time that I was selecting my research topic. In this conversation, I asked him whether Motorola would like me to focus in any particular topic. He then answered me that I had the freedom of

choosing the research topic and that Motorola understands that the value of its investment in my education is not in the outcome of a particular research project, but rather in the knowledge and skill that I was gaining while performing such research. Motorola has changed in these many years, as any company does and must do; however, its core values, which are every day expressed by its employees, have not changed, which makes Motorola a great company to work for. In particular, I thank the following individuals who made my Ph.D. journey possible: Jeff Bonta, Mark Birchler, Frank Yester, Dave Borth, Bob O'Dea, Faisal Ishtiaq, Dragan Boscovic, and Andy Aftelak. I also thank my various colleagues for their understanding of my time limitations. In particular, I thank my colleague and friend George Calcev for his support, for his interest, for his insightful questions, for his advice, and for his words of encouragement while I was tackling my most difficult research problems.

I also thank my wife, Patricia, for her continuous support since the day that I decided to apply to graduate school. Her support never wavered, even when we realized that attending graduate school would require me to work part-time and reduce our family income. I am lucky for having a wife who values higher education as much as I do. I also thank my son, Gustavo, who was born right after I started my Ph.D. journey, for the many hours of play that allowed me to relax my mind.

Finally, I thank my parents for providing me with the internal strength required for this and other accomplishments. Their teachings have made me believe in the pursuit of my goals and have shown me that they can be achieved with hard work and persistence. I still remember my father reciting verses of Rudyard Kipling's poem "If": *"If you can trust yourself when all men doubt you,"* and *"If you can force your heart and nerve and sinew. To serve your turn long after they are gone, and so hold on when there is nothing in you except the Will which says to them: Hold on!"* Especially this time, while pursuing a difficult goal through an unconventional path, their teachings kept me going.

## TABLE OF CONTENTS

	Page
<b>LIST OF FIGURES</b> . . . . .	viii
<b>LIST OF TABLES</b> . . . . .	x
<b>NOMENCLATURE</b> . . . . .	xi
<b>ABSTRACT</b> . . . . .	xiv
<b>1 Introduction</b> . . . . .	1
1.1 Motivation . . . . .	1
1.2 Issues Associated with the Analytical Treatment of Sensor Detection Systems . . . . .	4
1.3 Thesis Statement . . . . .	5
1.4 Overview of the Main Results . . . . .	6
1.5 Dissertation Outline . . . . .	7
<b>2 The Sensor Detection System Model</b> . . . . .	9
<b>3 Relevant Prior Research in Sensor Detection Systems</b> . . . . .	11
3.1 Conditionally Independent Measurements and Simple Hypotheses . . . . .	11
3.1.1 Unconstrained Communication Subsystems . . . . .	12
3.1.2 Sensors Making Local Decisions or Quantizing the Measurement . . . . .	13
3.1.3 Impairments of the Communication Subsystem . . . . .	17
3.1.4 Sensor Detection Systems Under Energy Constraints . . . . .	19
3.1.5 Sensor Detection Systems Over Multiple Access Communication Channels . . . . .	20
3.2 Conditionally Dependent Measurements . . . . .	22
3.2.1 Conditional Dependency Caused by Signal Randomness . . . . .	24
3.3 Composite Hypothesis . . . . .	27
3.3.1 Locally Optimum Approaches . . . . .	28
3.3.2 Approaches Based on a Least Favorable Distribution . . . . .	29

	Page
3.4 Summary . . . . .	30
<b>4 Problem Statement and Research Questions . . . . .</b>	<b>33</b>
4.1 Problem 1: Conditionally Dependent Measurements . . . . .	33
4.2 Problem 2: Composite Hypothesis . . . . .	34
4.3 Research Questions . . . . .	36
<b>5 Main Assumptions and Considerations . . . . .</b>	<b>37</b>
<b>6 Scenarios for Conditionally Independent and Conditionally I.I.D. Measurements . . . . .</b>	<b>40</b>
6.1 Emitter Location Distributed on a Specific Subset of the Region of Interest . . . . .	40
6.1.1 Example 6.1: . . . . .	45
6.2 Vanishing Amplitude Functions . . . . .	47
6.3 Sensors Deployed on a Sphere . . . . .	48
<b>7 Least and Most Favorable Distributions for Emitter Location . . . . .</b>	<b>50</b>
7.1 The Least Favorable Distribution for the Emitter Location . . . . .	51
7.1.1 The Least Favorable Distribution for a Class of Detection Systems . . . . .	52
7.1.2 The Least Favorable Distribution for a Detection System . . . . .	53
7.1.3 Maximin Detection Systems . . . . .	53
7.2 The Most Favorable Distribution for a Detection System . . . . .	55
7.2.1 Maximax Detection Systems . . . . .	56
7.3 Sufficient Conditions for Least and Most Favorable Distributions when Sensors Locations are Random Variables . . . . .	57
7.3.1 Example 7.1: . . . . .	69
7.3.2 Example 7.2: . . . . .	71
7.4 Summary . . . . .	73
<b>8 Using a Most Favorable Emitter Location Distribution to Evaluate How Conservative the Design Based on a Least Favorable Emitter Location Distribution is . . . . .</b>	<b>75</b>
8.1 Example . . . . .	77

	Page
8.2 Large Sample Methods to Evaluate How Conservative the Least Favorable Distribution is . . . . .	78
8.2.1 The ARE of the Maximax Centralized Detection System with Respect to the Maximin Centralized Detection System . . . . .	81
8.2.2 The ARE of Practical Centralized Detection Systems . . . . .	82
8.2.3 Bounds for the ARE Metric . . . . .	84
8.2.4 Numerical Evaluations . . . . .	87
8.2.5 Asymptotic Bounds for the ARE Metric . . . . .	88
8.2.6 The ARE Metric of a Maximin Centralized Detection System and a Practical Centralized Detection System Under a Least Favorable Distribution . . . . .	89
8.2.7 The ARE Metric of a Practical Centralized Detection System and a Practical Distributed Detection System Under a Least Favorable Distribution . . . . .	91
8.2.8 Relationship to Prior Research . . . . .	95
8.2.9 Summary of Analysis . . . . .	95
<b>9 Larger Deployment Regions . . . . .</b>	<b>97</b>
9.1 Motivation for Larger Sensor Deployment Regions . . . . .	97
9.2 Comparing Deployment Options using the ARE Metric . . . . .	99
9.3 Sufficient Conditions to Adopt Deployment Option 2 . . . . .	100
9.4 Numerical Evaluations . . . . .	103
9.5 Upper Bound for the Performance Gain of Deployment Option 2 . . . . .	105
9.6 The Impact of the Deployment Option 2 in the Performance Under other Distributions for the Emitter Location . . . . .	106
9.6.1 Example . . . . .	107
9.7 Summary . . . . .	108
<b>10 Summary of Contributions . . . . .</b>	<b>109</b>
<b>11 Conclusions and Future Research . . . . .</b>	<b>112</b>
11.1 Future Research . . . . .	113
<b>APPENDICES</b>	
Appendix A: Conditions under which Considering a Single Emitter is the Worst Case Scenario . . . . .	116



	Page
Appendix B: Scenarios in which Sensor Locations can be Considered Random Variables . . . . .	118
Appendix C: Supporting Lemmas for Chapter 6 . . . . .	119
Appendix D: Example that Illustrates the Need for an Augmented Definition for a Least Favorable Dis- tribution . . . . .	125
Appendix E: Considerations About the Conditional Distribution of the Measurements . . . . .	127
Appendix F: Supporting Propositions for Chapter 7 . . . . .	129
Appendix G: Supporting Example . . . . .	139
Appendix H: Detailed Calculations for the Example in Section 8.1 . . . . .	140
Appendix I: Supporting Material for Chapter 8 . . . . .	142
Appendix J: Computing the ARE Metric . . . . .	157
Appendix K: Supporting Propositions for Chapter 9 . . . . .	160
<b>LIST OF REFERENCES . . . . .</b>	<b>164</b>

## LIST OF FIGURES

Figure	Page
1.1 Sensor detection system with a fusion center in which each sensor $i$ collects the measurements $Z_i$ from a possible signal from an emitter at a random location, processes it with a sensor function $\phi_i$ , and transmits the sensor output $U_i$ to the fusion center, which applies a fusion function $\phi_0$ to decide between $H_0$ (signal absent) or $H_1$ (signal present). . . . .	1
2.1 Model for the sensor detection system being considered in this dissertation. . . . .	9
6.1 Non-regular convex hexagonal region of interest $S_e$ in which the distribution for the emitter location that places the emitter in each of the vertices of the polygon with probability $1/6$ does not cause the measurements $\{Z_i\}_{i=1}^K$ to become conditionally i.i.d. . . . .	45
7.1 Region of interest $S_e$ given by the union of a square and two triangles in which the distribution for the emitter location that places the emitter at $l_e^2$ is not always a least favorable distribution for a detection system. . . . .	70
7.2 Probability of detection under 0.01 probability of false alarm achieved by the centralized detection system $\phi_0^*$ under the distribution $P_{L_e}^{(r)}$ , where $P_{L_e}^{(r)}$ satisfies $P[L_e = (0, r)] = 1$ , for various values $r \leq r_M$ and for various ratios $r_A/r_M$ , where $r_A$ defines the function $\xi$ in in Example 7.2. . . . .	72
7.3 Conditional probability density functions $f_{Z_1 \theta=0}(z)$ (dashed and dotted line), $f_{Z_1 \theta=1, R=r_M}(z)$ (full line), $f_{Z_1 \theta=1, R=0}(z)$ (dashed line), and $H_1$ decision interval (shaded region) for 0.01 probability of false alarm when $r_A/r_M = 0$ (left), $r_A/r_M = 0.5$ (center), and $r_A/r_M = 0.75$ (right) in Example 7.2. . . . .	73
7.4 Probability of detection under 0.01 probability of false alarm achieved by the optimum centralized detection system $\phi_0^{**}$ under the distribution $P_{L_e}^{(r)}$ , where $P_{L_e}^{(r)}$ satisfies $P[L_e = (0, r)] = 1$ , for various values $r \leq r_M$ and for various ratios $r_A/r_M$ , where $r_A$ defines the function $\xi$ in Example 7.2. . . . .	74

Figure	Page
8.1 Probability of detection of the distributed detection system defined by (8.5) and (8.6) with the distribution of $Y_0$ and $k$ defined such that the probability of false alarm equals $\alpha_{\max} = 0.05$ when $L_e$ is distributed according to $P_{L_e}^-$ that places the emitter in the boundary of a circular region of interest with probability 1, $P_{L_e}^+$ that satisfies $P[L_e = (0,0)] = 1$ , and the uniform distribution $P_{L_e}^u$ for various values for the number of sensors $K$ . In all the evaluations, $\lambda_W = 1$ , $t_s = \lambda_W$ , $r_M = 15$ , and $A_{\max} = 40$ . . . . .	78
8.2 ARE of designs based on the most favorable distribution with respect to designs based on the least favorable distribution when $\xi(r) = 1/(1+r)^\kappa$ and $S_e = B_{r_M}(0)$ (top left), $S_e$ is an equilateral triangle (top center), and $S_e$ is a square (top right); and when $\xi(r) = \exp(-\kappa \cdot r)$ and $S_e = B_{r_M}(0)$ (bottom left), $S_e$ is an equilateral triangle (bottom center), and $S_e$ is a square (bottom right) for various values of $\kappa$ . . . . .	87
9.1 Two options for the sensor deployment region: $S_s^{(1)} = S_e$ (left) and $S_s^{(2)} = cS_e$ with $c = 1 + r_0/r_M$ (right). The shaded region represents the region in which $A_1 = 1$ when the emitter is located at $(0, -r_M)$ . . . . .	98
9.2 $ARE_{21}$ when $\xi(r) = 1/(1+r)$ (top), $\xi(r) = 1/(1+r)^2$ (middle), and $\xi(r) = \exp(-r)$ (bottom) for various values of $r_M$ when $S_e = B_{r_M}(0)$ (left), $S_e$ is a triangular region (middle), and $S_e$ is a square region (right). . . . .	104
9.3 Probability of detection of the distributed detection system defined by (8.5) and (8.6) with the distribution of $Y_0$ and $k$ defined such that the probability of false alarm equals $\alpha_{\max} = 0.05$ when $L_e$ is distributed according to $P_{L_e}^-$ that places the emitter in the boundary of a circular region of interest with probability 1, $P_{L_e}^+$ that satisfies $P[L_e = (0,0)] = 1$ , and the uniform distribution $P_{L_e}^u$ ; and when $\{L_i\}_{i=1}^K$ is uniformly distributed in $S_s = cS_e$ for either $c = 1$ (deployment option 1) or $c = 1.1$ (deployment option 2) for various values for the number of sensors $K$ . . . . .	107

LIST OF TABLES

Table	Page
D.1 Randomization parameter and decision threshold for the Neyman–Pearson detectors to satisfy a probability of false alarm equal to 0.05, probability of detection of the Neyman–Pearson detectors, and probability of detection when using the detector $\phi_{2/3}$ for each $c$ . . . . .	125

## NOMENCLATURE

$\mathbf{1}\{statement\}$	The indicator function: $\mathbf{1}\{statement\} = 1$ if <i>statement</i> is true, and $\mathbf{1}\{statement\} = 0$ if <i>statement</i> is false.
$P_X$	The probability measure induced by a random variable $X$ ; i.e., for any event $B$ , $P_X[B] := P[X \in B]$ .
$F_X$	The cumulative distribution function associated with $P_X$ ; i.e., $F_X(x) = P[X \leq x]$ ;
$P_{X C}$	For any random variable $X$ and any event $C$ , $P_{X C}$ represents the conditional probability measure of $X$ given $C$ ; i.e., $P_{X C}[B] := P[X \in B C]$ .
$\int g(x) dx$	The Riemann integral of $g(x)$ .
$\int g(x) dP_X(x)$	The Lebesgue integral of $g(x)$ with respect to the measure $P_X$ . <sup>1</sup>
$m(\mathcal{B})$	Lebesgue measure of a measurable set $\mathcal{B}$ .
$\ l\ $	The Euclidean norm in the space of $\delta$ dimensions; i.e., for any $l \in \mathbb{R}^\delta$ , $\ l\  := \sqrt{l'l}$ , where $l'$ is the transpose of the column vector $l$ .
$B_r(l)$	The closed ball of radius $r$ centered at $l$ ; i.e., $B_r(l) := \{x : \ x - l\  \leq r\}$ .
$\partial B_r(l)$	The boundary of the closed ball of radius $r$ centered at $l$ ; i.e., $\partial B_r(l) := \{x : \ x - l\  = r\}$ .
$B_r^\circ(l)$	The open ball of radius $r$ centered at $l$ ; i.e., $B_r^\circ(l) := \{x : \ x - l\  < r\}$ .
$H_0$	Null hypothesis in the decision problem: signal absent;
$H_1$	Alternative hypothesis in the decision problem: signal being emitted by an emitter at some point of the region of interest.

---

<sup>1</sup>For readers who are not familiar with Lebesgue integrals, it is possible to interpret the results of this dissertation with the usual Riemann integral if one considers that the random variable or vector under integration has a probability density function. For example, for any random  $X$  and any function  $g(X)$ , let  $f_X(x)$  be the probability density function of  $X$  and replace  $\int g(x) dP_X(x)$  with  $\int g(x) f_X(x) dx$ .

$\phi_0$	Fusion function used by the fusion center;
$\phi_i$	Sensor function used by the $i$ th sensor;
$\underline{\phi}$	Sensor detection system: $\underline{\phi} := \{\phi_i, P_{Y_i}\}_{i=0}^K$ .
$\alpha(\underline{\phi})$	Probability of false alarm obtained by $\underline{\phi}$ .
$\beta(P_{L_e}, \underline{\phi})$	Probability of detection obtained by $\underline{\phi}$ when $L_e$ is distributed according to $P_{L_e}$ .
$\xi$	Function that determines the amplitude of the signal component.
$\mathbf{dist}(L_i, L_e)$	Function that represents the distance between the sensor position $L_i$ and the emitter position $L_e$ .

### Commonly Used Variables and Sets

$K$	Number of sensors in the detection system.
$\theta$	Random variable that equals 0 when signal is absent and equals $\theta_1 > 0$ when the emitter generates the signal.
$L_e$	Random variable denoting the emitter location.
$S_e$	Region of interest; space in which $L_e$ takes value.
$L_i$	Random variable denoting the location of the $i$ th sensor.
$S_s$	Sensor deployment region; space in which $L_i$ takes value.
$Z_i$	Random variable denoting the measurement of the $i$ th sensor.
$A_i$	Random variable denoting the signal component of the measurement of the $i$ th sensor.
$W_i$	Random variable denoting the noise component of the measurement of the $i$ th sensor.
$U_i$	Random variable denoting the output of the $i$ th sensor function.
$Y_0$	Random variable used for randomization of the fusion function.
$Y_i$	Random variable used for randomization of the $i$ th sensor function.
$I_{0,y}$	Half-open or half-closed interval that determines the decision region in the fusion center when the randomization random variable $Y_0$ equals $y$ .

$I_{i,y,u}$	Interval of the form $(a,b)$ , $[a,b)$ , $(a,b]$ , or $[a,b]$ that determines the region in which the sensor function $\phi_i$ decides for the output $u$ when the randomization random variable $Y_i$ equals $y$ .
$t_{i,y,u}$	Threshold that defines the interval $I_{i,y,u}$ .

### Abbreviations

ARE	Asymptotic Relative Efficiency.
c.d.f.	Cumulative distribution function.
i.i.d.	Independent and identically distributed.
p.d.f.	Probability density function.
p.m.f.	Probability mass function.
SNR	Signal-to-noise ratio.
UMP	Uniformly Most Powerful.

## ABSTRACT

Sensor detection systems with a fusion center are being considered to detect a low-power signal emitter in an unknown location within a region of interest. For example, multiple radiation sensors can be distributed in a region to detect the presence of an unauthorized radioactive material. When designing such a system, the designer faces the problem that the measurements are conditionally dependent in general. Designs for conditionally dependent measurements are significantly more difficult to achieve than designs for conditionally independent measurements. Furthermore, the distribution of measurements depends on variables with unknown distributions, such as the emitter location, which means that the alternative hypothesis is composite. Although it is possible to use the theory of least favorable distributions to deal with a composite hypothesis, many of the results in this theory require conditionally independent measurements. If the designer assumes that measurements are conditionally independent, then any performance analysis may be invalid because such an assumption is generally considered unrealistic and only justified by convenience.

How can the designer deal with the difficulties associated with conditionally dependent measurements and the composite hypothesis? It is shown in this dissertation that there are conditions that allow a designer to deal with these problems by assuming a least favorable distribution for the emitter location that not only makes the hypothesis simple and ensures detection performance, but also causes the measurements to become conditionally independent and identically distributed. It is shown that under certain conditions, any distribution that places the emitter on a subset of the boundary of the region of interest with probability one is least favorable for various systems of interest. This result may be considered intuitive; however, it does not hold in general.

Since a design based on a least favorable distribution may be considered too conservative, this dissertation proposes the use of a most favorable distribution for the emitter location and uses the theory of asymptotic relative efficiency (ARE) to evaluate how conservative the design based on a least favorable distribution is. The ARE theory is further used to, under a least favorable distribution, compare different systems and different sensor deployment strategies.

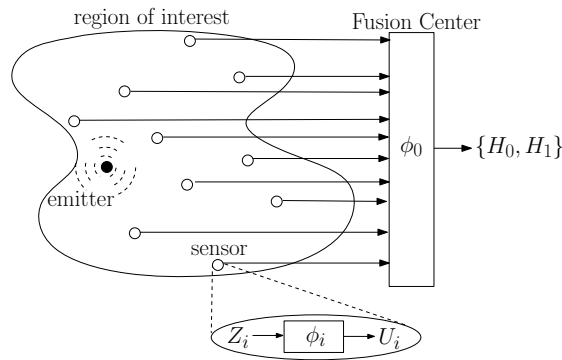


## Chapter 1

### Introduction

#### 1.1 Motivation

Consider the problem of **designing a sensor detection system with a fusion center to detect a low-power signal emitter** in an unknown location within a region of interest. As illustrated in Figure 1.1, a sensor detection system with a fusion center is a system in which multiple sensors at different locations collect measurements from the region of interest, process them, and transmit the processed outputs to the fusion center [117, 122]. Based on the information received from the sensors, the fusion center decides between the null hypothesis  $H_0$  (signal absent) and the alternative hypothesis  $H_1$  (signal being emitted by an emitter at some point of the region of interest.).



**Figure 1.1.** Sensor detection system with a fusion center in which each sensor  $i$  collects the measurements  $Z_i$  from a possible signal from an emitter at a random location, processes it with a sensor function  $\phi_i$ , and transmits the sensor output  $U_i$  to the fusion center, which applies a fusion function  $\phi_0$  to decide between  $H_0$  (signal absent) or  $H_1$  (signal present).

A sensor detection system with a fusion center may be used in many applications:

- **Detection of low-level point radiation sources:** It is desirable to detect as soon as possible a terrorism scenario in which an ionizing radioactive material such as Cesium-137 is released in a metropolitan area [20, 58, 87, 98]. The solution of placing individual portal-style detectors is not viable given the many possible entry points [87].

A sensor detection system in which sensors are radiation counters offers a solution to monitor and detect the presence of such point radiation sources in the whole region of interest [20, 58, 87, 98].

- **Detection of unused radio frequency communication channels:** A possible regulatory measure to increase the utilization of a radio frequency spectrum is to allow unlicensed radio devices to use a frequency band that is temporarily unused by its licensed users. Cooperative spectrum-sensing procedures are being proposed to determine whether a licensed user is transmitting in the frequency band and avoid excessive interference to licensed receivers [94, 111, 126]. Such schemes can be modeled by a sensor detection system in which sensors are radio frequency receivers.
- **Detection of forest fires:** It is desirable to detect forest fires as soon as possible in order to expedite the deployment of fire fighters to contain the spread of the fire. Satellite technology suffers from long scanning cycles, low resolution, and possible cloud cover [57, 103, 104, 140]. Sensor detection systems equipped with air temperature and humidity sensors are being considered as an option to detect the beginnings of a forest fire [57, 103, 104, 140].
- **Detection of intruders in a restricted area:** The movement of vehicles and persons causes vibrations in the ground that can be measured by seismic sensors [2, 93]. A sensor detection system with seismic sensors can detect intruders in an area and activate further systems such as an alarm or a video camera system [93].
- **Detection of unauthorized submarines in a region of the ocean:** Sonar systems for detecting submarines have been in use for many years [33, 66]; however, recent advances in submarine technology to avoid detection require new technologies to improve detection capabilities [6]. A sensor detection system with acoustic sensors can detect underwater sound waves generated or reflected by a submarine or a diver [6, 66, 105].
- **Detection of an aircraft in an air space:** Typical radar systems determine whether a region of air space contains an aircraft by using one or more stations to transmit a radio wave and one or more stations to receive the reflections that occur when an aircraft is present in the air space [7]. Modern radar systems are being considered in which information from multiple radar receivers is combined before reaching a decision [7, 76, 122, 134]. A sensor detection system in which sensors are radars and the signal emitted is the aircraft radio wave reflection can be used to model such modern radar systems [7, 67, 122, 134].

The sensor detection system in these applications may be a subsystem of a larger system that performs additional tasks such as emitter classification, localization, or tracking. In some larger systems, the task of detection is usually the first step in a sequence of tasks to deal with the emitter [24, 93, 125, 134]; for instance, in an intruder detection application, the video camera system is activated only after a seismic sensor detects the intruder [93]. In spite of the importance of the subsequent tasks, this dissertation deals exclusively with the design of the detection subsystem.

Sensor detection systems are being considered for the above mentioned applications because they enable measurements to be taken at closer distances from the potential emitter [43, 93]. In all of the applications mentioned, the measurements are corrupted by background noise and the amplitude of the signal to be detected decays as the distance from the emitter increases, and since the possible emitter is at an unknown location within the region of interest, collecting measurements from a single location can potentially produce measurements with very low signal-to-noise ratios (SNR). By distributing sensors at various locations, detection performance is improved in many situations [43, 93].<sup>1</sup> A further motivation for distributing the sensors is to allow the measurement of a possible emitter to be taken from different directions, reducing the impact of obstructions [43, 93].

Furthermore, sensor detection systems allow multiple measurements to be collected simultaneously, reducing the time required for detection. All of the above mentioned applications share the common goal of detecting the presence of the emitter with the smallest delay possible. Therefore, instead of having a single sensor sequentially collect multiple measurements to increase the SNR for decision, multiple sensors can each collect a single measurement and transmit it to the fusion center for a final decision in a fraction of the time.

From a practical point of view, there is an interest in using sensor detection systems because the communication of measurements from the sensors to the fusion center has been facilitated by the recent advances in wireless communication networks and protocols [2, 33, 93]. If communication cables were used to provide connectivity between sensors and fusion center, the system could not be economically viable due to the deployment costs; furthermore, deployment could be prevented altogether if sensors have to be deployed in a restricted or hard to reach region. By using wireless communication channels, deployment costs can be significantly reduced, making the system viable, and sensors can be deployed in a remote location; for instance, sensors can be dropped from an aircraft or contained within projectiles that scatter the sensors in the region [2, 57, 58, 125]. Advances in wireless ad hoc networking allow the self-organization of sensors in wireless ad hoc networks, which have been made possible by recent radio communication regulations and standards that use low-power radios in unlicensed radio frequency bands [2, 33]. The small cost of radios operating in such bands makes sensor systems with hundreds or thousands of sensors economically viable.

Lastly, sensor detection systems are more robust to failures and less prone to spoofing [2, 33]. The performance of a sensor detection system suffers slight degradation if a single sensor fails or is destroyed.

It is possible to classify the various types of sensor detection systems in two classes: systems that allow sensors to communicate with each other, and systems in which sensors communicate only with a fusion center or with a node at a higher level in a hierarchy than the sensors. Sensor systems that allow sensors to communicate with each other not only enable the formation of multihop communication paths towards the fusion center, but also enable the use of collaborative signal processing techniques in which a sensor uses both its own measurement and another sensor's measurement

---

<sup>1</sup>Distributing sensors at various locations will not always improve detection performance. There are contrived situations in which a better detection performance is obtained by placing all sensors at the center of the region of interest.

in the generation of its output [2, 33, 43]. On the other hand, such systems require more elaborate communication protocols and mechanisms for message routing, data aggregation, energy conservation, and self-configuration [2, 33].

In this dissertation, the focus is on sensor detection systems in which sensors communicate only with the fusion center, as illustrated in Figure 1.1. This is not limiting because such a sensor detection system can be considered one out of many fusion regions of a large hierarchical sensor detection system [44, 64, 136, 140]. In a hierarchical sensor detection system, a larger geographical area is partitioned in multiple fusion regions and the sensors in each region communicate with a single fusion node, which is supposed to decide between  $H_0$  and  $H_1$  within its fusion region.<sup>2</sup> This focus also does not preclude the case in which sensors communicate with the fusion center through data connections that use reliable transport protocols in a wired or wireless data network [71].

The design of sensor detection systems in which sensors communicate only with the fusion center is significantly less complicated because of fewer issues associated with the communication between sensors and fusion center and fewer issues associated with the control of sensor energy since sensors no longer relay transmissions from other sensors; however, the design of the sensor detection system is still complicated due to issues associated with the statistical treatment of measurements, as explained next.

## 1.2 Issues Associated with the Analytical Treatment of Sensor Detection Systems

The process of designing a sensor detection system with a fusion center involves the determination of how the sensors will process their measurements, how the sensors will transmit their measurements to the fusion center, how the fusion center will process the sensors' outputs, and how many sensors are needed in order to satisfy a prescribed performance requirement.

To produce such a design and evaluate its performance, the designer generally builds a mathematical model of the system and uses analytical treatment, numerical evaluations, simulation tools, and field tests to evaluate its performance. The analytical treatment allows the designer to better understand the underlying physical phenomena and provides guidance for the design; numerical evaluations and simulation tools are then used to refine the analysis; and field tests are used to verify the correct system operation in certain test conditions. This is a general procedure when designing any system and, ideally, the system designer should be equipped with good numerical and simulation tools, should have resources to perform field tests, and should have mathematical models that are amenable to analytical treatment.

This dissertation focuses on the analytical treatment of sensor detection systems, and **the problem that this dissertation addresses is that the current mathematical models for sensor detection systems are not amenable to analytical treatment unless the system designer adopts assumptions that are generally considered unrealistic by many researchers in the field.**

---

<sup>2</sup>In the literature, these fusion regions may be called 'clusters' and the fusion node may be called a 'clusterhead' or a cluster node [2, 44, 64, 136].

In detection problems, analytical treatment is facilitated when the data is assumed conditionally independent given the hypotheses [65]; in fact, under this assumption, models amenable to analytical treatment of sensor detection systems are available and much progress have been made, as described in Chapter 3; however, the assumption of conditional independence is generally considered unrealistic and only justified by convenience [15, 39, 106, 111, 124, 134, 137, 138]. As shown in Chapter 4, sensor measurements are indeed conditionally dependent in general when the distribution of the measurements depends on the distances between the sensors and the emitter locations [11, 39, 106]. With conditionally dependent measurements, the design becomes significantly more difficult even in simple sensor detection systems, as discussed in Chapter 3.

Furthermore, analytical treatment is generally difficult when the sensor measurements depend on variables with unknown distributions, which means that the hypothesis  $H_1$  is composite [65]. As discussed in Chapter 4, even if the distributions of some of the dependent variables can be accurately estimated or imposed, there are other variables that cannot be estimated or controlled in important situations. One such variable is the emitter location, whose distribution is generally unknown.

A valuable approach to deal with unknown distributions is to assume a worst case or *least favorable distribution* for the unknown distribution [73]. As explained in Chapter 7, a design that satisfies a certain detection requirement under the least favorable distribution for a system will also satisfy the detection requirement under any other distribution. However, as discussed in Chapter 3, many of the results in the theory of least favorable distributions require conditionally independent or i.i.d. measurements.

If the assumptions used in the analytical treatment are considered unrealistic, then the value of the analytical treatment for the design process is questionable.

The issues associated with the analytical treatment of sensor detection systems have been recognized by the research community. As put in [56], “*The reality is that for many practical applications, the mathematical assumptions upon which many of these methods are formulated are rarely satisfied. Sophisticated algorithms are easily corrupted and produce very poor results when the input data do not meet the requisite conditions (e.g., conditionally dependent observations, incorrect a priori probabilities, etc.)*”.

The difficulties in applying analytical treatment may discourage its use, and the system designer loses its potential benefits: overall guidance for the design, determination of optimal designs, determination of performance bounds, and comparison of candidate designs.

### 1.3 Thesis Statement

This dissertation is motivated by the following question: How can the system designer deal with the difficulties associated with conditionally dependent measurements and the composite hypothesis  $H_1$ ? If practical and realistic

ways to handle such difficulties can be found, then the analytical treatment becomes more amenable and appealing, and the system designer can defend its use.

This dissertation contains original evidence to support the following thesis statement:

**When detecting the presence of an emitted signal that decays with the distance between randomly distributed emitter and sensors, there are certain non-artificial conditions that allow a system designer to deal with the problems associated with the conditional dependence of measurements and the composite hypothesis by assuming a least favorable distribution for the emitter location that not only makes the hypothesis simple and ensures a detection performance, but also causes the measurements to become conditionally i.i.d., making models more amenable, allowing performance evaluations, and facilitating the use of large sample methods for the comparison of candidate designs.**

## 1.4 Overview of the Main Results

The results in this dissertation identify several sets of conditions that allow a system designer to avoid the complications associated with conditionally dependent measurements and a composite hypothesis. All of the results assume that both the emitter location is a random variable within a region of interest, and the sensors' locations are random variables in a deployment region that contains the region of interest.

A first set of results shows that, under certain scenarios and conditions on the distributions of emitter's and sensors' locations, the complications of conditionally dependent measurements are not present, and the measurements are actually conditionally i.i.d. This set of results indicates that, in the identified scenarios and conditions, the assumption of conditionally i.i.d. measurements is not unreasonable as argued by many authors.

A second set of results addresses the complications associated with a composite hypothesis through the theory of least favorable distributions. Although the theory of least favorable distributions is not new, its application to the distribution of the emitter location is new to the best of the author's knowledge.

Among these results, the main result is the identification of a least favorable emitter location distribution for a class of detection systems, and the conclusion that such a distribution is one of the distributions that cause the measurements to become conditionally i.i.d. This result indicates that, if the conditions are satisfied and a system designer adopts the identified least favorable emitter location distribution, then the designer will be able to not only operate with a simple hypothesis and ensure a detection performance, but also avoid the problem of conditional dependence among the measurements. The identified least favorable emitter location distribution places the emitter on a subset of the boundary of the region of interest with probability one. This result may appear intuitive and trivial; however, it is shown that such a distribution is not least favorable in all cases.

A third set of results illustrates how the theory of asymptotic relative efficiency (ARE) can be used to compare designs under a least favorable emitter location distribution.

Among these results, the asymptotic detection performance of a design under a least favorable emitter location distribution is compared against the asymptotic performance of a design under a *most favorable* emitter location distribution. These results allow a system designer to evaluate how conservative the design based on a least favorable emitter location distribution can be as the number of sensors grows.

The theory of ARE is also used to, under a least favorable emitter location distribution, compare different sensor detection systems. More specifically, centralized sensor detection systems; i.e., systems in which the fusion center has access to all the measurements, are compared against distributed sensor detection systems in which each sensor maps its measurement into a finite set.

The theory of ARE is further used to, under a least favorable emitter location distribution, compare sensor deployment strategies. It is shown that there are conditions under which it is better to deploy sensors in an enlarged region.

A detailed list of contributions is given in Chapter 10.

## 1.5 Dissertation Outline

Chapter 2 presents the model of the sensor detection systems considered in this dissertation. The goals of this section are: to present the key variables that define the sensor detection system, to precisely define the scope of this dissertation, and to provide a common model to describe the results achieved by other authors.

Chapter 3 presents a summary of the relevant results previously obtained by other authors, and illustrates that, although many results can be achieved under conditionally independent measurements and simple hypotheses, the design and analysis of sensor detection systems become significantly more complicated under conditional dependence and composite hypothesis; and fewer results are available under these conditions.

Chapter 4 contains this dissertation's problem statement and research questions.

Chapter 5 contains the main assumptions used in many of the results present in this dissertation.

Chapter 6 presents the first set of results, which cover the scenarios and conditions under which measurements are conditionally i.i.d. given either hypothesis.

Chapter 7 defines the concepts of least and most favorable distributions and applies them to the distribution of the emitter location. A least and a most favorable emitter location distributions for two classes of detection systems are identified in this chapter.

Chapter 8 compares the performance of several sensor detection systems designed under either a least or a most favorable emitter location distribution. The comparisons are performed using the ARE metric, which is explained in detail in Section 8.2.

Chapter 9 presents the conditions under which the deployment of sensors in an enlarged region can provide better detection performance.

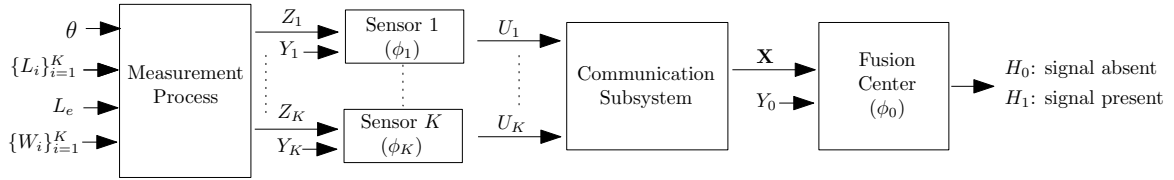
Chapter 10 contains a detailed list of contributions and the conclusions achieved from such contributions are in Chapter 11, which also contains a list of possible future research topics.



## Chapter 2

### The Sensor Detection System Model

Figure 2.1 illustrates the main components of the sensor detection system being considered in this dissertation: the measurement process,  $K$  sensors, the communication subsystem, and the fusion center.



**Figure 2.1.** Model for the sensor detection system being considered in this dissertation.

The measurement process determines the distribution of  $\mathbf{Z} := (Z_1, \dots, Z_K)$ , where  $Z_i$  represents the measurement obtained by the  $i$ th sensor. It is assumed that  $Z_i$  takes values on a set  $\mathcal{Z} \subset \mathbb{R}$ . Throughout this dissertation, the following additive model is assumed:

$$Z_i = \mathbf{1}\{\theta > 0\} \cdot A_i + W_i, \quad (2.1)$$

where  $\mathbf{1}\{statement\}$  equals 1 if the statement is true and 0 otherwise,  $\theta \in \{0, \theta_1\}$  is a random variable that equals  $\theta_1 > 0$  when the emitter generates the signal,  $W_i$  represents the noise in the measurement process, and  $A_i$  is a random variable that reflects how the signal is received by the sensor. The distribution of  $A_i$  depends on  $\theta$ , and on the distance between the sensor location  $L_i$  and the emitter location  $L_e$  through an amplitude function  $\xi$ . The sensors are located in the sensor deployment region  $S_s$ ; i.e.,  $L_i \in S_s$ ; and, when present, the emitter is located in the region of interest  $S_e$ ; i.e.,  $L_e \in S_e$ .

The measurements  $Z_i$  from each sensor  $i$  are transformed by the sensor function  $\phi_i$ . In order to allow randomization, the sensor function  $\phi_i$  also accepts as input a discrete random variable  $Y_i$  that takes values in  $\mathcal{Y}$ . In its most general form,  $\phi_i : \mathbb{R} \times \mathcal{Y} \rightarrow \mathcal{U}$  for some output space  $\mathcal{U}$ , and the sensor output is denoted by  $U_i := \phi_i(Z_i, Y_i)$ . It is assumed that a sensor  $i$  does not communicate with any other sensor or the fusion center to generate its output  $U_i$ .

The output  $U_i$  is subsequently transmitted to the fusion center through a communication subsystem, which outputs a random vector  $\mathbf{X}$ . Note that the model allows the possibility of a multiple access communication channel. If

the communication subsystem offers  $K$  individual and noninterfering communications channels, each of them used exclusively by one sensor, which means that  $\mathbf{X} := (X_1, \dots, X_K)$ , then it is called a *communication subsystem with dedicated channels*.

The output  $\mathbf{X}$  of the communication subsystem serves as input to the fusion function  $\phi_0$ . Based on  $\mathbf{X}$  and a possible randomization random variable  $Y_0$ , the fusion center makes the final decision between the hypotheses  $H_0 : \theta = 0$  and  $H_1 : \theta = \theta_1$ . In some portions of this dissertation, a sequence of hypotheses  $\{H_{1,n}\}_{n=1}^{\infty}$  in which  $H_{1,n} : \theta = \theta_{1,n} > 0$  with  $\theta_{1,n} \rightarrow 0$  is considered.

Certain types of sensor and fusion functions define the class of the detection system:

- If the sensor functions are such that  $\phi_i(z, y) = z$ , then the system is a *centralized detection system*.
- If the sensor functions are such that  $\phi_i : \mathcal{Z} \times \mathcal{Y} \rightarrow \mathcal{U}$  for any finite set  $\mathcal{U}$ , and the cardinality of  $\mathcal{U}$  is lower than the cardinality of  $\mathcal{Z}$ , then the system is a *distributed detection system*. If  $\mathcal{U}$  contains only two elements, then the system is a *binary distributed detection system*.
- If the sensor and fusion functions are such that, for any  $i \in \{0, 1, \dots, K\}$ ,  $\phi_i(z, y)$  remains constant as  $y$  varies, then the system is a *deterministic detection system*; otherwise, the system is a *randomized detection system*.

The acquisition and processing of each measurement  $Z_i$  determine the decision time interval of the detection system. Throughout the dissertation, it is assumed that  $\theta$ ,  $\{L_i\}_{i=1}^K$ , and  $L_e$  do not change during the decision time interval. Furthermore, motivated by the need to detect the presence of the emitter as soon as possible, this dissertation focuses on procedures in which the fusion center reaches the decision based on only the sensor measurements taken during a single decision interval; i.e., sequential decision procedures are not considered here.

The task of designing the sensor detection system can then be described as: given the characterization of the measurement process and the constraints on the communication subsystem, determine the number of sensors  $K$  and choose the sensor functions  $\{\phi_i\}_{i=1}^K$ , the fusion function  $\phi_0$ , and the distributions of  $\{Y_i\}_{i=0}^K$  in order to attain a prescribed detection performance.

This dissertation adopts the Neymann–Pearson framework, in which the detection performance requirement is specified as a minimum probability of detection under a maximum probability of false alarm [65]. The Bayesian framework, in which the system is designed to minimize the expected risk of a decision given a cost function and an *a priori* distribution for  $\theta$ , is not considered here.

## Chapter 3

### Relevant Prior Research in Sensor Detection Systems

This section begins with the description of previous results that confirm that the mathematical models for sensor detection systems are amenable when deciding between two simple hypothesis and the measurements are either conditionally independent or conditionally i.i.d. Under these conditions, much progress have been made in both centralized and distributed sensor detection systems under both the Neyman–Pearson or Bayesian frameworks.

Relevant results that attempt to deal with conditionally dependent measurements and composite hypotheses are described next. It will become clear to the reader that designs under such conditions are significantly more complicated.

#### 3.1 Conditionally Independent Measurements and Simple Hypotheses

Assuming that measurements are conditionally independent or conditionally i.i.d. given  $\theta$ , and assuming that both  $H_0$  and  $H_1$  are simple hypotheses,<sup>1</sup> many researchers have successfully modeled and analyzed sensor detection systems from various points of view:

- **Unconstrained communication subsystems:** The sensor measurements  $\{Z_i\}_{i=1}^K$  are transmitted directly to the fusion center, forming a centralized detection system;
- **Sensors making local decisions or quantizing the measurement:** The sensor outputs  $\{U_i\}_{i=1}^K$  are constrained to belong to a set  $\mathcal{U}$  with finite cardinality, forming a distributed detection system;
- **Impairments in the communication subsystem:** The communication subsystem introduces noise or fading in the received signal at the fusion center and sensors have a maximum transmit power constraint such that sensor transmissions may be received with errors;
- **Energy Constraints:** Sensors are battery-powered, which motivates techniques to minimize unnecessary sensor operation and conserve energy;

---

<sup>1</sup>A hypothesis is simple if the measurements under the hypothesis can be characterized by a single distribution; otherwise, the hypothesis is composite [65].

- **Multiple Access Communication Channels:** Sensors no longer have dedicated communication channels to transmit to the fusion center.

### 3.1.1 Unconstrained Communication Subsystems

When the communication subsystem does not impose any constraints, the sensors can simply collect their measurements and forward them to the fusion center, forming a centralized detection system. In this case, the design task becomes the determination of the fusion function  $\phi_0$ , its randomization random variable  $Y_0$ , and the number of sensors  $K$  to achieve a prescribed detection performance.

If the conditional distribution of the measurements  $\{Z_i\}_{i=1}^K$  under  $H_0$  and the conditional distribution of  $\{Z_i\}_{i=1}^K$  under  $H_1$  are completely characterized, it is possible to find the optimum design using classical detection theory [65]. The pair  $(\phi_0, Y_0)$  that maximizes the probability of detection for a prescribed probability of false alarm is given by the Neyman–Pearson Lemma [65], and the minimum number of sensors  $K$  required to attain a given performance specification can be determined, particularly if the sensor measurements are conditionally independent or conditionally i.i.d. given the hypothesis.

The design and analysis of the centralized detection system under simple hypotheses  $H_0$  and  $H_1$ , and under conditional independence of measurements can therefore be accomplished without many challenges. Some specific examples can be found in [94, 111, 138]. The authors in [94, 111] studied the centralized detection system in a collaborative spectrum-sensing system in which the sensor measurements form a multivariate Gaussian vector under either hypothesis with independent components. A similar model was developed by the authors in [138], which considered the centralized detection system as a benchmark for other detection systems.

When measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. and both  $H_0$  and  $H_1$  are simple hypotheses, it is also possible to analyze centralized detection systems using traditional large sample methods.

A large sample method used by many authors is the asymptotic decay exponent. In this method, the performance metric is given by the asymptotic decay exponent of the probability of miss under a constraint on the probability of false alarm. The asymptotic decay exponent for a centralized detection system when both  $H_0$  and  $H_1$  are simple hypotheses and the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. can be obtained from Stein’s Lemma [70], which says that

$$\lim_{K \rightarrow \infty} \frac{1}{K} \log P[\phi_0^{(K)}(\mathbf{Z}, Y_0^{(K)}) = 0 | \theta = \theta_1] = -D(P_{Z_i|\theta=0} \| P_{Z_i|\theta=\theta_1}), \quad (3.1)$$

where  $\{\phi_0^{(K)}\}_{K=1}^\infty$  and  $\{Y_0^{(K)}\}_{K=1}^\infty$  are respectively the Neyman–Pearson detectors and associated randomization variables that satisfy a common constraint in the probability of false alarm for each  $K$ ,  $D(P_{Z_i|\theta=0} \| P_{Z_i|\theta=\theta_1})$  is the Kullback–Leibler divergence between the distribution of  $Z_i$  conditioned on  $\theta = 0$  and the distribution of  $Z_i$  conditioned on  $\theta = \theta_1$ .

One of the motivations for focusing on the asymptotic decay exponent is that it allows the estimation of an approximate number of sensors required to achieve a given detection performance. Some authors [81, 82] used the asymptotic decay rate of the centralized detection system as a bound for other detection systems.

Another possible large sample method considers the Asymptotic Relative Efficiency (ARE) metric between two candidate detection systems [73]. As explained in Section 8.2, these methods consider how fast the number of sensors in each candidate system must grow as the signal amplitude decays to zero in order to ensure that the probabilities of detection and false alarm converge to a desired performance in each candidate detection system. As will be seen in the next section, this approach has been used to compare centralized and distributed detection systems.

### 3.1.2 Sensors Making Local Decisions or Quantizing the Measurement

If the sensor outputs  $\{U_i\}_{i=1}^K$  are constrained to belong to a set  $\mathcal{U}$  with finite cardinality, then the system becomes a distributed detection system. In such a system, the design task becomes the determination of the set of fusion and sensor functions  $\{\phi_i\}_{i=0}^K$ , their randomization random variables  $\{Y_i\}_{i=0}^K$ , and the number of sensors  $K$  to achieve a prescribed detection performance.

This section describes relevant prior results to enable the design of the distributed detection system when sensor functions are given and just the fusion function needs to be designed, when both sensor and fusion functions must be designed, and when the number of sensors is asymptotically large. All of these results assume simple hypotheses and conditionally independent measurements.

#### 3.1.2.1 Optimal Fusion Function Design for Distributed Detection Systems

Consider the task of determining the optimal fusion function for a distributed detection system that uses given binary sensor functions; i.e.,  $\mathcal{U} = \{0, 1\}$ . Assume a communication system that offers dedicated and error-free channels such that the fusion center receives  $\{U_i\}_{i=1}^K$ .

In this case, classical detection theory can be used to determine the optimal fusion function in either the Neyman–Pearson or Bayesian frameworks. In either framework, the optimal fusion function uses the weighted sum of sensor outputs  $U_i$  as decision statistic, in which the weight for each  $U_i$  is determined by its distribution [21]. The case in which the system uses identical sensor functions is highlighted:

**Theorem 3.1.** *In a distributed detection system with given binary sensor functions in which the sensor outputs  $\{U_i\}_{i=1}^K$  are conditionally i.i.d. with known distributions under  $H_0 : \theta = 0$  and under  $H_1 : \theta = \theta_1$ , and the communication system that offers dedicated and error-free channels, the fusion function  $\phi_0$  that maximizes the probability of detection for a given probability of false alarm is given by*

$$\phi_0(\mathbf{u}, y_0) = \mathbf{1} \left\{ \sum_{i=1}^K u_i \in I_{0, y_0} \right\}, \quad (3.2)$$

where  $I_{0,1} = (k, \infty)$ ,  $I_{0,2} = [k, \infty)$ ,  $y_0$  is the realization of the randomization random variable  $Y_0 \in \{1, 2\}$ , which has Bernoulli distribution and is independent of  $\theta$ . Both  $k$  and the parameter of the Bernoulli distribution are chosen to satisfy the constraint on the probability of false alarm.

Theorem 3.1 is the Chair and Varshney rule for identical sensors [21], also known as the  $k$ -out-of- $K$  rule, in the Neyman–Pearson context with a randomized threshold.

### 3.1.2.2 Optimal Sensor and Fusion Function Design for Distributed Detection Systems

If the sensor functions  $\{\phi_i\}_{i=1}^K$  and their respective randomization random variables  $\{Y_i\}$  are to be defined together with the fusion function, then the design of an optimal set of sensor and fusion functions is significantly more complicated, even if the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. with known distributions under  $\theta = 0$  and  $\theta = \theta_1$ .

Regarding the optimization of sensor functions  $\{\phi_i\}_{i=1}^K$ , it was shown by many [13, 21, 112, 117, 131] in different degrees of generality that: if the sensor measurements  $\{Z_i\}_{i=1}^K$  are conditionally independent, if  $H_0$  and  $H_1$  are simple hypotheses, and if each sensor has a dedicated error-free communication channel with the fusion center, then, under either the Bayesian or Neyman–Pearson decision frameworks,

- The optimal sensor function  $\phi_i$  of each sensor is a likelihood ratio quantizer [112, 117];
- The optimal decision regions used by a sensor function  $\phi_i$  depend on the decision regions used by the fusion function and other sensor functions [112];
- The optimal sensor function of a sensor  $i$  may be different from the optimal sensor function of a sensor  $j$ , even when the observations  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. [112].

The form of the optimal sensor and fusion functions that maximize the probability of detection for a given constraint on the probability of false alarm is given by the following theorem:

**Theorem 3.2.** [114, 117] *Consider a distributed detection system with a communication subsystem offering dedicated and error-free communication channels. Let  $y$  be the realization of a random variable  $Y \in \{1, \dots, M\}$  with given distribution  $P_Y$  independent of  $\mathbf{Z}$  and  $\theta$ . Assume that  $P_{Z_i|\theta=\theta_1}$  is absolutely continuous with respect to  $P_{Z_i|\theta=0}$ , which is absolutely continuous with respect to a common measure  $\mu$ . Let  $f_{Z_i|\theta=\theta_1}(\mathbf{z})$  and  $f_{Z_i|\theta=0}(\mathbf{z})$  be the corresponding density functions. If the measurements  $\{Z_i\}_{i=1}^K$  are conditionally independent, then*

- The optimal fusion function is given by

$$\phi_0(\mathbf{u}, y) := \mathbf{1} \left\{ \prod_{i=1}^K \mathcal{L}_{U_i, S'_e, y}(u_i) \in I_{0,y} \right\}, \quad (3.3)$$

where  $\{I_{0,y}\}_{y=1}^M$  are intervals defined by  $(t_{0,y}, \infty]$  for certain values  $\{t_{0,y}\}_{y=1}^M$ ; and

$$\mathcal{L}_{U_i, S'_e, y}(u_i) := \begin{cases} \frac{P_{U_i|\theta=\theta_1, Y=y}(u_i)}{P_{U_i|\theta=0, Y=y}(u_i)}, & P_{U_i|\theta=0, Y=y}(u_i) > 0, \\ +\infty, & P_{U_i|\theta=0, Y=y}(u_i) = 0. \end{cases} \quad (3.4)$$

- The optimal sensor functions  $\phi_i(z_i, y)$  satisfy

$$P[\phi_i(Z_i, Y) = u, \mathcal{L}_Z(Z_i) \notin I_{i,Y,u} | \theta = \theta^*] = 0 \quad (3.5)$$

for either  $\theta^* = 0$  or  $\theta^* = \theta_1$ , where

$$\mathcal{L}_Z(z_i) := \begin{cases} \frac{f_{Z_i|\theta=\theta_1}(z_i)}{f_{Z_i|\theta=0}(z_i)}, & f_{Z_i|\theta=0}(z_i) > 0, \\ +\infty, & f_{Z_i|\theta=0}(z_i) = 0, \end{cases} \quad (3.6)$$

and, for each  $i, y$ , and  $u \in \{0, \dots, U^{\max}\}$ ,  $I_{i,y,u}$  is an interval defined by thresholds  $t_{i,y,u}$  and  $t_{i,y,u+1}$ , and such thresholds must satisfy

$$0 = t_{i,y,0} \leq t_{i,y,1} \leq \dots \leq t_{i,y,U^{\max}} \leq t_{i,y,U^{\max}+1} = \infty \quad (3.7)$$

for a given integer  $U^{\max} > 0$  imposed by the communication subsystem.

Theorem 3.2 summarizes results from [114] and [117]. In more details, it was proven in [114] that the sensor and fusion functions that maximize the probability of detection given the constraint that the probability of false alarm be equal or below  $\alpha_{\max}$  are such that the sensor functions satisfy condition (3.5) with  $y$  being realizations of independent random variables  $\{Y_i\}_{i=1}^K$ ; and the fusion function uses the likelihood ratio of the sensor outputs as its decision statistic. Such a distributed detection system is said to have *independent randomization*. It was proven in [117] that a distributed detection system that maximizes the probability of detection and has probability of false alarm equal to  $\alpha_{\max}$  exists and randomizes between two sets of fusion and sensor functions that have independent randomization. It is also proven in [117] that the performance of the optimal distributed detection system with independent randomization can be achieved by a distributed detection system with fusion and sensor functions as defined in Theorem 3.2 with random variables  $Y_1 = \dots = Y_K = Y$  almost surely. Such a distributed detection system is said to have *dependent randomization*.

Optimal sensor and fusion functions for the Bayesian framework also follow the form of (3.3) and satisfy (3.5); however, they do not require the randomization of sensor and fusion functions. This indicates that the optimal design under the Neyman–Pearson framework are somewhat more difficult to find than the optimal design under the Bayesian framework.

Although it is possible to write the form of the optimal sensor and fusion functions, it is difficult to analytically determine the various intervals  $\{I_{0,y}\}_{y=1}^M$  and  $\{\{I_{i,y,u}\}_{u=0}^{U^{\max}}\}_{y=1}^M\}_{i=1}^K$  because such intervals need to be *jointly optimized*. The joint optimization is required because the definition of the intervals  $\{\{I_{i,y,u}\}_{u=0}^{U^{\max}}\}_{y=1}^M$  used in sensor  $i$  depends on the definition of the intervals used in other sensors and on the fusion function intervals  $\{I_{0,y}\}_{y=1}^M$  [112]; furthermore,

the intervals  $\{\{I_{i,y,u}\}_{u=0}^{U^{\max}}\}_{y=1}^M$  used in sensor  $i$  may be different from the intervals used by any other sensor, even under conditionally i.i.d. measurements [112].

The search for the optimal decision intervals can be significantly simplified under conditions identified by the authors in [133]. Under a general setting that includes both the conditionally dependent and conditionally independent cases, the authors in [133] showed that, under conditions that include binary sensor functions and positive finite likelihood ratios with no point masses under either hypothesis, the optimum sensor functions for a given fusion function under the Neyman–Pearson framework would all use a common test threshold.

Numerical procedures to optimize sensor and fusion functions were proposed [122, 132]; however, they do not guarantee that the optimal set of sensor and fusion functions will be found [122, p. 75] [24, 124].

### 3.1.2.3 Large Sample Methods to Simplify the Design of Distributed Detection Systems

Given the difficulty in reaching optimal designs, many researchers have used large sample methods to reach designs that are asymptotically optimal as the number of sensors increase. Large sample methods are approximate; however, they are widely accepted when properly used [73, 121].

Many of the authors that considered large sample methods focused on the asymptotic decay exponent of the probability of miss under a constraint in the probability of false alarm [25, 116].

Using asymptotic decay exponents, the author in [116] showed that, under simple hypotheses  $H_0$  and  $H_1$  and considering a dedicated and error-free communication subsystem, if sensor measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. under either  $H_0$  or  $H_1$ , then the distributed detection system with the highest asymptotic decay exponent has all sensors using a common sensor function that quantizes the likelihood ratio of  $Z_i$  and maximizes  $D(P_{U_i|\theta=\theta_0} \| P_{U_i|\theta=\theta_1})$ ; thus, when considering a system with a large number of sensors, a system designer may consider a common sensor function and the design would be significantly simplified.

Also considering the asymptotic decay exponent as the performance metric, the authors in [22] compared systems with binary sensors and systems with sensors quantizing at a higher number of bits. For this comparison, the authors in [22] considered a constraint in the total number of bits that can be reliably transmitted over the communication subsystem and considered how to best allocate the bits to the sensors. Under this setting and considering conditionally i.i.d. observations, simple hypotheses, and the Bayesian framework, the authors in [22] showed that there are conditions under which the highest decay exponent is achieved when using binary sensors; i.e., it is better to have more sensors providing observations quantized into a single bit than having less sensors providing observations quantized into multiple bits. The rationale for this result is that there are conditions under which the first bit of an observation provides more than half of the Chernoff information in the unquantized observation [109]. The authors showed that the conditions are met when the noise random variables  $\{W_i\}_{i=1}^K$  are Gaussian or exponential distributions.



Other authors considered large sample methods based on the Asymptotic Relative Efficiency (ARE) metric or similar metrics [37, 128, 129] to study certain types of distributed detection system.

The authors in [37] claimed an analytic expression for the ARE of a centralized detection system with respect to a binary distributed detection system in which all sensors compare the likelihood ratio against a common threshold. The obtained expression for the ARE depends on the probability density function (p.d.f.) of the additive noise  $W$ . For the Gaussian p.d.f. with zero mean and finite variance, the authors claimed that the ARE metric equals  $\pi/2$  and, for the Laplacian density p.d.f. with finite variance, the authors claimed that the ARE metric equals 1, claiming that the distributed and centralized detection systems are asymptotically equivalent in this case.

Focusing on the distributed detection system that uses binary sensor functions of the form  $\phi_i(z_i, y_i) = \mathbf{1}\{z_i > t\}$ , the authors in [128, 129] computed the efficacy of such a class of detection systems. Considering that the additive noise has a p.d.f.  $f_W(w)$ , the authors in [128, 129] showed that the efficacy is proportional to  $f_W(0)^2$  when  $t = 0$ ; and concluded that, if  $f_W(0) = 0$ , such a distributed detection system would become ‘useless’ because the resulting ARE of any other detection system (with a non-zero efficacy) with respect to this distributed detection system would be zero.

### 3.1.3 Impairments of the Communication Subsystem

When focusing on the possible impairments of the communication subsystem, several authors considered that dedicated, interference-free, but error-prone communication channels would be available for each of the sensors. More precisely, using the model of Chapter 2, each sensor transmits its sensor output  $U_i \in \mathcal{U}$  over a dedicated communication channel  $(\mathcal{U}, f_{X|U=u}(x), \mathcal{X})$ , whose output defines the random variable  $X_i \in \mathcal{X}$ , and the outputs  $\{X_i\}_{i=1}^K$  of the various communication channels are used by the fusion center for decision.

Many authors [31, 45, 99, 113] considered simplified models for the communication channel, adopting binary symmetric channels to represent the possible communication errors in a distributed detection system. Other authors [26, 27, 62, 79, 80, 84, 89] considered more elaborate models to study the effect of wireless communication channels. The relevant results under both types of models are described in the next subsections.

#### 3.1.3.1 Binary Symmetric Channels

When the communication subsystem offers  $K$  dedicated and non-interfering binary symmetric channels, it is possible to extend the Theorem 3.1 and determine the optimal fusion function for a binary distributed detection system with given sensor functions. Considering the Bayesian framework with conditionally independent measurements, the authors of [31] investigated the form of the optimal fusion function for given identical binary sensor functions transmitting over binary symmetric channels with identical crossover probabilities. The authors concluded that the optimal

fusion function is still in the form of the  $k$ -out-of- $K$  sensors and it is possible to extend the authors' result to the Neymann–Pearson framework.

Regarding optimal fusion and sensor functions in the Neyman–Pearson framework, the authors in [99] and [113] showed that the optimal sensor functions for a given fusion function are still based upon the likelihood ratio of the measurements. They further showed that the optimal decision threshold of a sensor is not only a function of the thresholds used in other sensors, but also a function of the crossover probabilities of other sensors' channel. The authors in [99] also provided some examples that suggest that significant performance loss may occur when the communication channel introduces errors but the sensor functions are designed assuming perfect communication channels.

The authors in [45] have considered the case when a subset of the  $K$  binary symmetric channels are error-free while the remaining channels have non-zero crossover probabilities, and evaluated two design options to fuse conditionally independent measurements: use all observations  $\{X_i\}_{i=1}^K$  or use just the subset of observations that arrive in error-free communication channels. Numerical results suggest that the later option is preferred when the measurement SNR is high or the crossover probability is high.

### 3.1.3.2 General Communication Subsystems with Dedicated Channels

The authors in [27, 79] have shown the optimality of likelihood ratio quantizers under the Bayesian framework when measurements  $\{Z_i\}_{i=1}^K$  are conditionally independent, the communication channel is modeled as  $K$  dedicated and interference-free communication channels, and the communication subsystem output space  $\mathcal{X}$  is an arbitrary vector space. In the same setting, the authors in [79] showed that optimum quantization levels may use less than the total number of bits available for quantization depending on the communication channel reliability. More reliable channels would quantize  $Z_i$  into more quantization levels than less reliable channels. When using less than the allowed number of bits for quantization, the remaining bits could be used to reduce the probability of reception error at the fusion center.

Focusing on wireless communication channels, the authors in [26, 62, 80, 84, 89] considered that each one of  $K$  dedicated channels would be impaired by additive white Gaussian noise (AWGN) and flat fading and, considering given sensor functions, investigated various candidate designs for the fusion function, each candidate applying strategies from the field of wireless communications to improve the reception of the sensor outputs under various assumptions of knowledge of the fading statistics.

There are also various studies that considered transmission power control algorithms for wireless sensor networks; see for instance [69, 91] and the references within. Such algorithms can reduce the energy consumption in sensors; however, the transmitter is not the sole consumer of energy in a sensor and other energy conserving techniques are required.

### 3.1.4 Sensor Detection Systems Under Energy Constraints

Given that many envisioned sensor detection systems would contain battery-powered sensors, many authors considered techniques to conserve sensor energy [91, 96, 102]. Many of them focused on the design of communication protocols and operation schemes to reduce the expected energy consumption of sensors [3, 32, 91, 96]; however, these studies did not consider the impact in the detection performance.

Among the studies that focused on the impact of energy constraints in the detection performance, the authors in [25] showed that sensor detection systems under power constraints can be evaluated through the augmentation of the commonly used Chernoff information (CI) or Kullback–Leibler (KL) divergence. Considering conditionally i.i.d. observations and simple hypotheses  $H_0$  and  $H_1$ , the authors proposed that the CI or KL divergence be normalized by the expected power consumed by each sensor. The authors further proposed that this metric be used to compare candidate sensor detection systems, and provided examples in which the metric was used to compare the performance of binary and analog sensors. Considering the Bayesian framework, their numerical examples suggest that the normalized CI is higher when using analog sensors and the measurement SNR is small; however, as the measurement SNR increases, the normalized CI of binary sensors becomes higher.

Other studies considered detection systems in which sensors do not always transmit to the fusion center in order to reduce the energy consumption. These systems are usually classified as systems having *censoring* mechanisms.

#### 3.1.4.1 Sensor Detection Systems with Censoring

In a sensor detection system with censoring, sensors are exempt to transmit when their observation falls within a region considered to be ‘less informative’; i.e., the sensor output space  $\mathcal{U}$  contains a ‘no-send’ element  $\Xi$  that satisfies  $P[U_i = \Xi | \theta = 0] > 0$  and  $P[U_i = \Xi | \theta = 1] > 0$  for at least some  $i \in \{1, \dots, K\}$ . This scheme was originally proposed and studied by [97] in order to reduce the expected energy consumed in each of the sensors.

A censoring distributed detection system can be seen as an ON-OFF signaling scheme in which sensors would not transmit to the fusion center when  $U_i = 0$ . With this point of view and considering faded wireless communication channels between binary sensors and the fusion center, the authors in [62] have considered the design of the fusion function based on the energy detected in each sensor transmission.

When designing a sensor detection system with censoring, it is usual to consider a constraint on the transmission rate of sensors [5, 97]. Such a constraint is equivalent to a set of constraints on the conditional distributions of the sensor output  $U_i$ . While considering the constraint

$$\sum_{i=1}^K P[U_i \neq \Xi | \theta = 0] \leq \gamma \quad (3.8)$$

for some constant  $\gamma \in (0, 1)$ , the authors of [97] and [5] focused on the design of optimal sensor functions and on the form of the region  $\{Z_i : \phi_i(Z_i) = \Xi\}$  that defines the set of observations in which each sensor  $i$  would not transmit.

The authors in [97] concluded that, under simple hypotheses  $H_0$  and  $H_1$  and either the Bayesian or the Neyman–Pearson frameworks, if the sensor measurements  $\{Z_i\}_{i=1}^K$  are conditionally independent, each sensor has a dedicated error-free communication channel to the fusion center, and each sensor adopts a deterministic decision function that either transmits the likelihood ratio of  $Z_i$  or does not transmit, then the optimal sensor decision function that satisfies (3.8) is such that  $U_i = \phi_i(Z_i) = \Xi$  if the likelihood of  $Z_i$  falls within a closed interval of the real line, where the bounds of the closed interval may vary from sensor to sensor.

The authors in [5] have extended this result by claiming that the optimal sensor function under the Neyman–Pearson framework would be a randomized function between three of the deterministic functions defined in [97]. Under the same assumptions, the authors in [5] have also shown that, if the sensor output space  $\mathcal{U}$  are restricted to be a finite set and under the constraint (3.8), then the optimal fusion and sensor functions uses dependent randomization, as defined in Section 3.1.2.2.

The authors in [17, 18] have proposed a procedure that indirectly produces a censoring centralized detection system through the medium access method. In such a method, each sensor obtains the likelihood ratio of its measurement and starts a timer that is inversely proportional to the magnitude of the log-likelihood ratio. When the timer expires, the sensor transmits the log-likelihood to the fusion center. Thus, sensors with more relevant measurements would transmit earlier to the fusion center. By imposing a maximum time duration for the transmissions in a given decision interval or a feedback mechanism, sensors with low values of log-likelihood ratio would be prevented to transmit and their energy would be conserved with little impact to the final decision.

### 3.1.5 Sensor Detection Systems Over Multiple Access Communication Channels

While all of the works mentioned so far assumes that the communication subsystem offers a dedicated communication channel for each of the  $K$  sensors, some authors considered that the  $K$  sensors are accessing a single multiple access communication channel with the fusion center. More precisely, the communication subsystem offers a single communication channel defined by  $(\mathcal{U}^K, f_{X|U=\mathbf{u}}(x), \mathcal{X}^L)$ , in which  $L \cdot |\mathcal{X}| < K \cdot |\mathcal{U}|$ , where  $|\mathcal{A}|$  represents the cardinality of a set  $\mathcal{A}$ .

The main motivation for considering multiple access communication channels is the observation that offering  $K$  dedicated and non-interfering communication channels, as considered in all the works mentioned in the previous subsections, would require the number of communication channels to grow at the same rate as the number of sensors, which could be difficult to satisfy if the number of sensors is large.

Since the fusion center does not necessarily require the individual sensor measurements to reach its decision; i.e., a fusion center may decide based upon a statistic of the sensor measurements and a correct decision may be reached even if no individual sensor measurement is received, it is not necessary to equip the communication subsystem with one dedicated communication channel for each sensor. By considering that all  $K$  sensors access a single multiple access

communication channel, it is possible to design a system in which  $L \cdot |\mathcal{X}|$  remains bounded while the number of sensors  $K$  grows.

In this setting, the authors in [41] have provided the following general result regarding the optimum sensor functions when considering  $|\mathcal{X}| < \infty$  and  $|\mathcal{U}| < \infty$ : in either the Bayesian or the Neyman–Pearson frameworks, if the sensor measurements  $\{Z_i\}_{i=1}^K$  are conditionally independent, then communication channel errors can be modeled as probability mass functions (p.m.f.s)  $P[X_1 = x_1, \dots, X_L = x_L | U_1 = u_1, \dots, U_K = u_K]$  and

- The optimal sensor function  $\phi_i$  of each sensor is a likelihood ratio quantizer;
- The definition of the likelihood ratio quantizer used in sensor  $i$  depends on the p.m.f.s  $P[X_1 = x_1, \dots, X_L = x_L | U_1 = u_1, \dots, U_K = u_K]$ , the definition of the likelihood ratio quantizer used in other sensors, and the definition of the fusion function.

Many authors have considered multiple access communication subsystems specifically designed for wireless sensor systems [52, 78, 81, 82, 135].

A simple but important example is the communication subsystem in which sensors modulate a common waveform to transmit to the fusion center, which receives the sum of the received signals [52, 74, 78, 81]. The authors in [81] showed that a sensor detection system under conditionally independent measurements using such a communication subsystem would have an exponential decay in the error probability. The authors in [78] derived the form of the optimal fusion and sensor functions under the Bayesian framework when using this communication subsystem with Rayleigh fading, showing that sensors that decide based upon the likelihood ratio of conditionally independent measurements are still optimal.

Receiving the sum of all transmissions not only is efficient in the usage of communication channels, but may also provide a better detection performance than the communication subsystem that offers  $K$  communication channels. For instance, in a Gaussian channel and under conditionally independent measurements, better detection performance is achieved because the fusion center operates with a statistic that has  $K$  times less variance in the noise introduced by the communication subsystem [81].

Motivated by this communication SNR gain, many authors have considered the performance of different sensor functions under the single-channel communication subsystem. A sensor function usually considered is the *amplify-and-forward*, in which  $U_i = Z_i$  and pulse amplitude modulation (PAM) is used to generate a common waveform with  $U_i$  as its amplitude. Another common sensor function is the *decode-and-forward*, in which the space of measurements is partitioned in two subsets,  $U_i \in \{-1, 1\}$ , and the sensor would use a binary phase shift keying (BPSK) modulator to transmit over the communication subsystem. Modified amplify-and-forward, modified decode-and-forward, and sensor functions that combine both strategies were studied by [74] and [52].

Another important multiple access scheme proposed for sensor detection systems is the type-based multiple access (TBMA) procedure. In this scheme, sensors cooperate to transmit the histogram of the measurements  $\{Z_i\}_{i=1}^K$  to the

fusion center [81]. The authors in [81] showed that a sensor detection system with such a communication subsystem and under conditionally independent measurements would have an exponential decay in the error probability. The authors of [82] reached similar results and also considered the impact of fading in the communication channel.

It is important to mention that the communication subsystems mentioned in this subsection require that sensors' transmissions be synchronized in order to be added coherently in the fusion center. Such a requirement has motivated research to enable the synchronization of sensors' transmissions [85, 86].

### 3.2 Conditionally Dependent Measurements

When the measurements  $\{Z_i\}_{i=1}^K$  are no longer conditionally independent, the system designer faces two major challenges to design and evaluate a sensor detection system:

- In general, it is difficult to characterize the joint distribution of the measurements  $\{Z_i\}_{i=1}^K$ ; and
- It is difficult to determine optimal designs when designing distributed detection systems;

The characterization of the joint distribution of several random variables that are not independent is difficult in general. To illustrate this point, consider the case in which the  $K$  measurements are Bernoulli random variables. Under independence, one only needs to know  $K$  parameters to characterize the joint distribution of the  $K$  random variables. Under dependence, one needs to know  $P[Z_j = 1 | Z_{j-1} = z_{j-1}, \dots, Z_1 = z_1]$  for all  $2^{j-1}$  possible values of  $Z_{j-1}, \dots, Z_1$  and for all  $j \in \{1, \dots, K\}$ , which results in a total of  $2^K - 1$  parameters.

In order to allow analytical treatment, many authors, particularly in the area of cooperative spectrum-sensing, assume that the measurements are distributed according to the multivariate Gaussian distribution [1, 23, 35, 36, 72, 75, 94, 111, 120, 126, 138]. In this case, the designer has the expression for the joint distribution of the measurements. Furthermore, the use of the multivariate Gaussian distribution allows the use of large sample theory results to determine the exponential decay rate of the probability of errors as the number of sensors grow to infinity [8, 23, 75]. In either type of analysis, it is necessary, however, to determine the vector of expected values and the covariance matrix. Usually, authors assume that the correlation between any two measurements is equal to a common correlation parameter [1, 111]; or assume that the correlation is a function of the distance between the two sensors [23, 75, 120, 126].

When considering the multivariate Gaussian distribution to design a centralized detection system, the optimal fusion function in either the Neyman–Pearson or Bayesian frameworks can be defined and the resulting performance can be evaluated since the likelihood ratio of the measurements can be treated analytically using classical detection theory [23, 75, 94, 111, 126, 138].

Considering the multivariate Gaussian distribution does not, however, necessarily allow the optimal design of distributed detection systems. In such systems, difficulties arise in the design of both the fusion and the sensor functions.

Regarding the fusion function, the optimal function is not as simple as in Theorem 3.1 [39,63]. Under the Bayesian framework and conditionally dependent measurements, the authors in [63] showed that the form of the optimal fusion function for a given set of binary sensor functions involves the multiplication of two terms: the first term is the likelihood ratio that would be obtained if the measurements were conditionally independent, and the second term is a function of the various correlation coefficients among the sensor outputs. The optimal fusion function design would require knowledge or estimation of  $2(2^K - K - 1)$  correlation coefficients [63]. Although one could obtain the correlation coefficients of the binary sensor outputs when sensors observe measurements distributed according to the multivariate Gaussian distribution and use local likelihood ratio detectors to define their binary output, this task is not trivial, as exemplified by [63] for a three-sensor system.

Regarding the sensor functions, one of the most important available results is that optimal sensor functions for either the Neyman–Pearson or the Bayesian frameworks may or may not be likelihood ratio quantizers; i.e., the region in which a binary sensor  $i$  decides for  $U_i = 1$  may not be a single interval in the range of the measurements' likelihood ratio as in the conditionally independent case [30, 88, 115, 117, 130]. To illustrate that this applies even in simple settings, consider a sensor detection system with two binary sensors under conditionally dependent measurements following a multivariate Gaussian distribution. Under the Bayesian framework, the authors in [30] and [130] have shown that there are sufficient conditions on the distribution parameters under which the likelihood ratio quantizer is optimal; however, there are other sufficient conditions on the distribution parameters under which the likelihood ratio quantizer is not optimal.

Since the region in which a sensor  $i$  decides for  $U_i = 1$  may contain multiple intervals in the range of the measurements' likelihood ratio, the numerical procedures to determine the optimal sensor functions need to search over a larger space of candidate sensor functions [115].

In some special cases, it is possible to reduce the region in which a sensor  $i$  decides for  $U_i = 1$  into a single interval under a different decision statistic [13, 133]. Under the Neyman–Pearson framework and certain conditions that include binary sensor functions and measurements with distributions that do not contain point masses, the author of [13] has shown that the optimum sensor function for a given fusion function decides  $U_i = 1$  when the ratio of a function that depends on both the likelihood of the measurement  $Z_i$  and the correlation between  $Z_i$  and  $Z_j$  for  $j \neq i$  is above a threshold. In a subsequent paper, the authors of [133] showed that, for a given fusion function, the optimal sensor thresholds would be common for all sensors.

Although the results in [13] and [133] simplify the search for the optimal sensor functions, having to know the correlation among measurements may still be difficult, particularly if the sensors' relative locations are not known during the design phase.

Large sample theory can be used to derive the form of the asymptotically optimal sensor function in some particular cases [23]; however, numerical procedures would still be required to determine the different coefficients that define the sensor function. The authors in [14] derived the form of the asymptotically optimal sensor function that maximizes

the efficacy of a locally optimum fusion function when the measurements are conditionally dependent following a dependency model in which  $Z_i$  and  $Z_j$  are independent if  $\|i - j\|$  exceeds an imposed integer value. Given that the resulting asymptotically optimal sensor function depends on scalar coefficients for which there are no closed form solutions, the authors in [14] provide an iterative algorithm to determine the final form of the sensor functions.

Given the difficulty in determining optimal sensor functions in either a finite or an asymptotic setting, several authors considered suboptimal distributed detection designs [1, 72, 77, 120]. Under the Bayesian framework, the authors in [72] found the best binary sensor function in the class of likelihood ratio tests when considering sensors observing conditionally dependent measurements that follow the multivariate Gaussian distribution, concluding that the determination of the optimal sensor thresholds for a two-sensor system would require the solution of coupled nonlinear equations. The authors in [1, 120] also considered binary sensors that decide based on the likelihood ratio of their observations and performed numerical experiments. In [1], it is shown that negative correlation in the measurements may increase detection performance while positive correlation may decrease detection performance when compared to the conditionally independent case.

A special case of conditionally dependent measurements occurs when the measurements  $\{Z_i\}_{i=1}^K$  are conditionally dependent given  $\theta$  but are conditionally independent given  $\theta$  and additional random variables. The authors in [28] and [29] have studied the design of optimum sensor functions in this scenario. The authors assume that the measurements are conditionally dependent given  $\theta$  but conditionally independent given a hidden random variable that depends on  $\theta$ . The introduction of the hidden random variable allows the designer to optimize the sensor function of a sensor  $i$  for a given set of sensor functions used by other sensors and a given fusion function. This result may facilitate the design of numerical procedures to search for optimal sensor functions under conditionally dependent measurements.

A further special case of conditionally dependent measurements, which is an instance of the special case in which  $\{Z_i\}_{i=1}^K$  are conditionally dependent given  $\theta$  but conditionally independent given  $\theta$  and additional random variables, is described in the next section.

### 3.2.1 Conditional Dependency Caused by Signal Randomness

An important special case is the one in which the measurements follow the additive model of (2.1); i.e., each measurement  $Z_i$  is the sum of a signal random variable  $A_i$  and a noise random variable  $W_i$ , and the noise random variables  $\{W_i\}_{i=1}^K$  are conditionally i.i.d. and independent of  $\{A_i\}_{i=1}^K$  and  $\theta$ . The randomness of the random variable  $A_i$  can cause the measurements  $\{Z_i\}_{i=1}^K$  to be conditionally dependent given  $\theta$ ; however,  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  and  $\{A_i\}_{i=1}^K$ ; i.e., the dependency of  $\{Z_i\}_{i=1}^K$  given  $\theta$  is induced by the dependency among  $\{A_i\}_{i=1}^K$ . Many authors provided results under this assumption [35, 38, 46, 90, 106, 119, 127, 134].

As previously mentioned, the multivariate Gaussian distribution can be used to model the distribution of the measurements  $\{Z_i\}_{i=1}^K$ , where the dependency among  $A_i$  is considered in the nondiagonal covariance matrix that defines the



alternative hypothesis  $H_1$ . This approach is commonly used in cooperative spectrum-sensing applications [111, 126]; however, it cannot be used in other important applications, such as detection of low-power radiation sources, in which the measurements are radiation counts.

In order to model the dependency of  $\{A_i\}_{i=1}^K$  when detecting a low-power radiation source, the authors in [106] used an approximation in which the joint distribution of the measurements is decomposed as the product of the marginal distributions and a particular function of the marginal distributions. With this approximation, the authors study a fusion function for a binary distributed detection system.

Another approach to model the dependency of  $\{A_i\}_{i=1}^K$  is to assume that they form a Markov Random Field [38, 127]. In this model, for any sensor  $i$ , the conditional distribution of  $A_i$  given the realization of all other  $A_j$ , for  $j \neq i$ , is equal to the conditional distribution of  $A_i$  given the realizations of only the  $A_j$  from sensors near sensor  $i$ . The authors in [38] proposed an iterative method to estimate the realizations of each  $A_i \in \{0, 1\}$  and determine whether a sensor is within an ‘event-region’; i.e., a region in which a particular condition is present. The authors in [127] also considered the event-region detection problem and proposed a sensor function that decides based upon its current measurement, its past decisions, and past decisions from sensors in its neighborhood. In both [38, 127], sensors exchange information with each other.

The authors in [35, 36] have proposed dealing with the dependency among the random variables  $\{A_i\}_{i=1}^K$  through a piecewise approximation. In this approximation, the random variables  $\{A_i\}_{i=1}^K$  are partitioned in groups. Within each group, the realizations of  $A_i$  are set equal to each other; and  $A_i$  and  $A_j$  from different groups are treated as independent. This approximation is motivated when the dependency among  $\{A_i\}_{i=1}^K$  is caused by the proximity of sensors: sensors located close to each other have similar realizations of  $A_i$ ; i.e., the values of  $A_i$  would be strongly correlated; and sensors far from each other have approximately independent values of  $A_i$ . The authors use this model to analyze a hierarchical sensor classification system in which the region of interest is partitioned into many subregions and, for each subregion, an intermediary node would collect measurements of sensors located within the subregion and send an average of the measurements to the fusion center. The sizes of the subregions are designed such that the measurements from sensors within the subregion can all be part of a single group in their measurement model. By assuming that all random variables  $A_i$  are fully correlated within a group and independent across groups, the design and analysis of sensor detection systems become significantly simpler, as shown in [35, 36] when considering measurements following the multivariate Gaussian distribution. Even when the multivariate Gaussian distribution is not a suitable model, this approximation simplifies the design and analysis of detection systems because the conditional distribution of measurements given  $\theta = \theta_1$  within a given group can be obtained through the integration of the conditional distribution of measurements given  $\theta$  and given all  $A_i$  of the group over the realizations of a single random variable  $A_i$  of the group.

### 3.2.1.1 Special Case: Signal with a Distribution that is a Function of $\xi(\|L_i - L_e\|)$

An important case is when the signal  $A_i$  has a distribution that depends on the distance between the sensor location  $L_i$  and an emitter location  $L_e$  through a nonincreasing function  $\xi$ . This case is being considered in this dissertation and has been considered by authors working in applications such as radar detection [134], detection of low-power radiation sources [98, 106], cooperative spectrum-sensing [94, 111, 126], and others [90, 119].

When the emitter or sensor locations are random variables, the analysis is difficult because, as shown in detail in Chapter 4, the random variables  $\{A_i\}_{i=1}^K$ , and therefore the measurements  $\{Z_i\}_{i=1}^K$ , are conditionally dependent in general, even when the distributions of the emitter and sensor locations are known.

In this case, one approach to avoid the complications of conditionally dependent measurements is to assume that the locations of the sensors and the emitter are given. In this case,  $\{A_i\}_{i=1}^K$  are no longer random and the measurements become conditionally independent. This approach can be justified in cooperative spectrum-sensing applications when the sensors are at fixed and known locations and the emitter is at a stationary and known location; e.g., a television broadcast station [94]; however, when the primary user of a spectrum is mobile, this assumption is harder to justify.

Still in cooperative spectrum-sensing applications, it is common to assume that the sensors are all located at approximately the same distance from the emitter [111, 120, 126]. In other words, the dependency among  $\{A_i\}_{i=1}^K$  is avoided by assuming that  $\{A_i\}_{i=1}^K$  are deterministic variables and equal to each other. The authors in [120] treated  $\{A_i\}_{i=1}^K$  as equal to each other by assuming that all the sensors would be located at the edge of the guard band of the primary user of the spectrum; i.e., the primary user would be located at the center of a disk and all sensors locations  $\{L_i\}_{i=1}^K$  are on the boundary of the disk.

When the emitter location is not available, some authors have assumed it to be an unknown but deterministic quantity [98, 119]. In [98], the authors propose that sensors at fixed and known locations initially estimate the emitter's location; then, using the estimated location as the actual emitter location, decide whether the emitter is present or not. With sensor and emitter locations given, the measurements become conditionally independent with known distributions. Considering that the emitter location is in fact a nuisance parameter, this approach can be viewed as an instance of the Generalized Likelihood Ratio Test (GLRT) [65]. The authors in [119] considered a binary distributed detection system to detect the presence of an emitter at an unknown but deterministic location  $l_e$  and proposed the use of the following performance metric:

$$\beta_{\min} := \min_{l_e \in S_e} \beta(l_e), \quad (3.9)$$

where  $\beta(l_e)$  is the probability of detection when the emitter is at location  $l_e$ . The authors argued that the lack of knowledge regarding the emitter location motivates the use of the  $k$ -out-of- $K$  fusion function and claimed that, with such a fusion function and under a conditionally i.i.d. noise process, sensor functions that use likelihood ratio detectors maximize the minimum probability of detection when one varies the emitter location. Considering the 'OR' fusion function

and assuming that all binary sensor functions use the same threshold, the authors studied sensor deployment strategies in a line segment and argued that placing sensors in regular intervals performs better than random deployment.

A related approach is to consider separate detection problems for each possible location for the emitter. In this case, the sensor detection system would perform multiple measurements and tests and the fusion center would reach a decision regarding the presence of an emitter in each of the possible locations of the emitter. This approach was taken by the authors in [134], which considered a centralized detection system in which sensors at deterministic locations are radars to detect the presence of a target within a region of interest. The region of interest is partitioned in several subregions and sensors probe each subregion one at a time, according to a schedule. Since the probing of each subregion is attempting to decide between the null hypothesis and the alternative hypothesis that there is a target at the location corresponding to the subregion, the random variable corresponding to the location of the target becomes deterministic for the decision problem and the signal random variables are no longer conditionally dependent.

The authors in [90] recognized the difficulty of designing a detection system when the emitter location is randomly distributed and, like [119], argued that the lack of knowledge regarding the emitter location motivates the use of the  $k$ -out-of- $K$  fusion function. They further imposed that the fusion center treat all sensor outputs  $\{U_i\}_{i=1}^K$  as conditionally independent. Assuming that the sensor locations  $\{L_i\}_{i=1}^K$  and the emitter location  $L_e$  are all uniformly distributed in the region of interest, they computed approximations for the detection performance of the system under the imposed conditional independence.

### 3.3 Composite Hypothesis

When the conditional distributions of the measurements  $\{Z_i\}_{i=1}^K$  are not completely characterized; i.e., if either  $H_0$  or  $H_1$  is a composite hypothesis, then the design becomes significantly more challenging for the following reasons:

- The Uniformly Most Powerful (UMP) detector; i.e., the fusion function that maximizes the probability of detection for a given probability of false alarm for any possible pair of distributions from  $H_0$  and  $H_1$ , may not exist [73, p. 83].
- Even if the system designer adopts a suboptimum fusion function and applies numerical methods to determine whether a design satisfies a prescribed performance, the designer would have to ensure that the prescribed performance is achieved for each possible pair of distributions from  $H_0$  and  $H_1$  and uncountable possible pairs may exist.

One way to deal with the composite hypothesis problem is to use the Generalized Likelihood Ratio Test (GLRT) and estimate the one or more parameters that cause  $H_1$  to be composite [65]. The GLRT approach was also explored in [107], in which the authors considered a practical binary distributed detection system and proposed the estimation of each of the unknown Bernoulli parameters that characterize the distribution of each sensor's output to build the fusion

function. The GLRT approach was also used by the authors in [98] while proposing the estimation of the emitter's location and using the estimate to build the decision test, as detailed in Section 3.2.1.1.

Two additional ways to deal with the composite hypothesis problem are described in the next subsections.

### 3.3.1 Locally Optimum Approaches

A type of composite hypothesis explored by many authors is the situation in which various distributions within the hypothesis  $H_1$  are parameterized by a scalar parameter  $\theta$ ; i.e.,  $H_0 : \theta = 0$  and  $H_1 : \theta > 0$ , and the typical approach in this case is to consider locally optimum detectors [9–12, 14–16, 95, 108, 128, 129]. In all these studies, the authors used a measurement model given by  $Z_i = \theta \cdot A_i + W_i$  in which  $A_i$  may be a deterministic variable or a random variable with known distribution that does not depend on  $\theta$  such that the distribution of  $Z_i$  is completely characterized for each possible value of  $\theta$ .

Among the various results obtained in this setting, the authors in [15] derived the form of the locally optimal sensor and fusion functions for a binary distributed detection system with either conditionally independent or conditionally dependent measurements. When the measurements are conditionally independent, the optimum sensor function has the same form as if the sensor were the single sensor in the system. When the measurements are conditionally dependent, the optimum sensor function has an additional term that depends on various joint conditional distributions among the fusion center output and other sensors' outputs [15, 16]. These results were extended to the non-binary distributed detection case in [14].

The authors in [108] proposed a locally optimum approach that uses randomly deployed binary sensors to detect a known spatial signal and approached the detection problem through the detection of a nonhomogeneous point process. Sensors are randomly deployed according to a homogeneous spatial Poisson process. Under  $H_0$ , sensors with  $U_i = 1$  form a homogeneous Poisson process; however, under  $H_1$ , they form a nonhomogeneous Poisson process because the spatial signal makes the intensity of the process location-dependent. The authors claim the asymptotically locally most powerful detector in the fusion center when sensor functions produce  $U_i = \mathbf{1}\{Z_i > t\}$ .

The authors in [139] applied the locally optimum approach to a cooperative spectrum-sensing application, treating the distributions of the received signals at the fusion center as Gaussian-distributed with known variances and treating the mean of such Gaussians as the unknown parameter that causes  $H_1$  to be composite.

It is important to observe that the locally optimum approach is actually a large sample method since its optimality, when the decision is based on conditionally i.i.d. measurements, is asymptotic as  $\theta_1 \rightarrow 0$  in an appropriate manner as the number of sensors grows to infinity [73, p. 546]. Such a large sample method is based on the concept of efficacy and the Asymptotic Relative Efficiency (ARE) metric between two candidate detection systems, which are described in detail in Section 8.2.

### 3.3.2 Approaches Based on a Least Favorable Distribution

A different approach to deal with the problem of composite hypothesis  $H_1$  is to build a simple hypothesis using the least favorable distribution within  $H_1$ . Assuming the least favorable distribution and the distribution under  $H_0$ , the system is designed using the techniques for the simple hypotheses case. The motivation for such an approach is to build a detection system under a worst-case condition such that the resulting design have a guaranteed detection performance under any of the other distributions within  $H_1$ .

Approaches based on least favorable distributions are also considered in this dissertation and the concept of a least favorable distribution is described in detail in Chapter 7; however, in order to present the results from other authors [4, 50, 59, 73, 123] in a precise manner, the following definitions are adopted for this section: let  $P_{Z_1, \dots, Z_K | \theta = j}$  denote the distribution of the measurements  $\{Z_i\}_{i=1}^K$  under  $H_j$  for  $j \in \{0, 1\}$ , let  $\beta(P_{Z_1, \dots, Z_K | \theta = 1}, \underline{\phi})$  and  $\alpha(P_{Z_1, \dots, Z_K | \theta = 0}, \underline{\phi})$  respectively denote the probability of detection obtained by the detection system  $\underline{\phi}$  under the distribution  $P_{Z_1, \dots, Z_K | \theta = 1}$  and the probability of false alarm obtained by  $\underline{\phi}$  under the distribution  $P_{Z_1, \dots, Z_K | \theta = 0}$ . For any pair of distributions  $P_{Z_1, \dots, Z_K | \theta = 1}$  and  $P_{Z_1, \dots, Z_K | \theta = 0}$  and a given desired maximum probability of false alarm  $\alpha_{\max}$ , let

$$\underline{\phi}_{\{P_{Z_1, \dots, Z_K | \theta = 0}, P_{Z_1, \dots, Z_K | \theta = 1}\}} := \underset{\phi: \alpha(P_{Z_1, \dots, Z_K | \theta = 0}, \phi) \leq \alpha_{\max}}{\operatorname{argmax}} \beta(P_{Z_1, \dots, Z_K | \theta = 1}, \underline{\phi}); \quad (3.10)$$

i.e., the sensor detection system  $\underline{\phi}_{\{P_{Z_1, \dots, Z_K | \theta = 0}, P_{Z_1, \dots, Z_K | \theta = 1}\}}$  is the Neyman–Pearson detection system when considering  $P_{Z_1, \dots, Z_K | \theta = 0}$  and  $P_{Z_1, \dots, Z_K | \theta = 1}$  as the distribution of  $\{Z_i\}_{i=1}^K$  under  $H_0$  and  $H_1$  respectively.

When the detector has access to all conditionally i.i.d. measurements; i.e., when the sensor detection system is centralized, Lehmann [73] defines that  $P_{Z_1, \dots, Z_K | \theta = 0}^*$  and  $P_{Z_1, \dots, Z_K | \theta = 1}^*$  form a pair of *least favorable distributions* if  $\forall P'_{Z_1, \dots, Z_K | \theta = 0}, P'_{Z_1, \dots, Z_K | \theta = 1}$ ,

$$\beta(P'_{Z_1, \dots, Z_K | \theta = 1}, \underline{\phi}_{\{P'_{Z_1, \dots, Z_K | \theta = 0}, P'_{Z_1, \dots, Z_K | \theta = 1}\}}) \geq \beta(P_{Z_1, \dots, Z_K | \theta = 1}^*, \underline{\phi}_{\{P_{Z_1, \dots, Z_K | \theta = 0}^*, P_{Z_1, \dots, Z_K | \theta = 1}^*\}}); \quad (3.11)$$

i.e., an optimal system for the pair of least favorable distributions has the lowest probability of detection for a given maximum probability of false alarm than the optimal system for any other pair of distributions. In Chapter 7, the limitations of such a definition are explored.

Also considering the case in which the detector has access to all conditionally i.i.d. measurements, Huber [59] assumes that the hypotheses  $H_0$  and  $H_1$  are each given by a nominal distribution, with distributions in its vicinity following an  $\varepsilon$ -contaminated model, and shows that there is a pair of distributions  $P_{Z_1, \dots, Z_K | \theta = 1}^-$  and  $P_{Z_1, \dots, Z_K | \theta = 0}^-$  such that, for any other pair  $P'_{Z_1, \dots, Z_K | \theta = 1}$  and  $P'_{Z_1, \dots, Z_K | \theta = 0}$ ,

$$\beta(P'_{Z_1, \dots, Z_K | \theta = 1}, \underline{\phi}_{\{P_{Z_1, \dots, Z_K | \theta = 0}^-, P_{Z_1, \dots, Z_K | \theta = 1}^-\}}) \geq \beta(P_{Z_1, \dots, Z_K | \theta = 1}^-, \underline{\phi}_{\{P_{Z_1, \dots, Z_K | \theta = 0}^-, P_{Z_1, \dots, Z_K | \theta = 1}^-\}}) \quad (3.12)$$

and  $P_{Z_1, \dots, Z_K | \theta = 1}^-$  and  $P_{Z_1, \dots, Z_K | \theta = 0}^-$  satisfies (3.11); i.e., it is a least favorable pair of distributions in the sense of Lehmann [73]. In [60], Huber and Strassen extended this conclusion to the case in which the distributions of each composite hypothesis are upper bounded by a common set function termed *2-alternating capacities*.

Assuming composite hypotheses  $H_0$  and  $H_1$  with conditionally i.i.d. measurements that satisfy (3.12), the authors in [50] concluded that  $P_{Z_1, \dots, Z_K | \theta=1}^-$  and  $P_{Z_1, \dots, Z_K | \theta=0}^-$  would also satisfy (3.12) if the Neyman–Pearson centralized detection system is replaced by a distributed detection system with an asymptotically large number of sensors using identical binary sensor functions that decide based on the likelihood ratio of the observations. The authors in [50] further concluded that the asymptotic exponential decay rate of the considered distributed detection system designed under the least favorable pair of distributions is a bound for the decay rates of the distribution detection system under any other distribution.

The authors in [123] extended the results of [50] for general distributed detection systems with a finite number of sensors, reaching the important conclusion that the pair of distributions  $P_{Z_1, \dots, Z_K | \theta=1}^-$  and  $P_{Z_1, \dots, Z_K | \theta=0}^-$  that satisfies (3.12), which considers a centralized detection system, is the same pair of distributions that satisfies (3.12) when considering an optimal distributed detection system under conditionally independent measurements.

Other authors also considered least favorable distribution approaches but in less related settings: the authors in [4] extended results of [123] to censored detection systems; and the authors in [53] adopted the Bayesian framework and considered least favorable distributions for the distribution of  $\theta$ .

### 3.4 Summary

When measurements are conditionally independent and the hypotheses  $H_0$  and  $H_1$  are simple, there is a large body of research that describes the form of optimal centralized and distributed sensor detection systems and analyzes their performance under various conditions, under energy constraints, and under various types of communication subsystems. When designing under conditionally independent measurements, a system designer can rely on the following key results:

- The forms of the optimal centralized and distributed detection systems are known for either the Bayesian and Neyman–Pearson frameworks. Optimal fusion and sensor functions use the likelihood ratio as decision statistic [21, 112, 117];
- Optimal distributed detection designs contain many peculiarities: the optimal decision regions used by a sensor function depend on the decision regions used by the fusion function and other sensor functions; and the optimal decision regions of a sensor may be different from the optimal decision regions of other sensors, even when the observations are conditionally i.i.d. [112]. Furthermore, optimal designs under the Neyman–Pearson framework generally require dependent randomization of sets of fusion and sensor functions [117]. Nevertheless, numerical local optimization procedures exist [122, 132];
- Large sample methods can be used to provide guidance for the design of systems with a large number of sensors. Asymptotic decay rates for distributed detection systems were obtained under various conditions and constraints

[22, 25, 116] and the Asymptotic Relative Efficiency (ARE) metric can be used to compare candidate design approaches [37, 128, 129];

- Most of the optimality results can be extended to the case in which the communication channels between sensors and fusion center are error-prone [27, 41, 79, 99, 113]. Performance results of various types of sensor detection systems that use specific wireless communication channels are available [26, 52, 62, 78, 80–82, 84, 89, 135]; and
- Censoring schemes are available to reduce the energy consumption in sensors [5, 17, 18, 62, 97].

When measurements are no longer conditionally independent, it is difficult to determine designs of optimal distributed detection systems [13, 30, 39, 63, 88, 115, 130, 133]. The difficulty stems not only from the complicated form of the optimal fusion and sensor functions, but also, and more importantly, from the lack of information about the correlation parameters. Although some special cases allow simplifications, approximations, and analytical treatment [35, 38, 46, 90, 106, 119, 127, 134], particularly if the distances between sensor and emitter locations are given [94, 111, 120, 126], system designers often have to settle for suboptimal designs [1, 35, 36, 38, 46, 72, 77, 90, 106, 119, 120, 127].

Furthermore, in general, it is difficult to characterize the joint distribution of the measurements  $\{Z_i\}_{i=1}^K$  when they are conditionally dependent. The most common approach to characterize the joint distribution and model the dependency among measurements is to assume the multivariate Gaussian distribution [1, 23, 35, 36, 72, 75, 94, 111, 120, 126, 138]. Although the adoption of the multivariate Gaussian distribution allows analytical treatment, it is still difficult to characterize the correlation among  $\{Z_i\}_{i=1}^K$ . In order to proceed, simple correlation models have to be assumed [1, 23, 75, 111, 120, 126].

In the special case in which the dependency among measurements is caused by randomness in the emitter's location, some authors proposed approaches that circumvent the conditional dependency by either requiring that sensors execute estimation procedures to initially estimate the emitter's location [98] or requiring that the sensors and fusion center perform multiple measurements and tests for each of the possible locations of the emitter [134].

When either of the hypotheses  $H_0$  or  $H_1$  is composite, there are at least three approaches to deal with the composite hypothesis problem:

- The generalized likelihood ratio test (GLRT), which involves the estimation of unknown quantities to build the likelihood test [98, 107];
- Locally optimum approaches, which assume that  $H_0 : \theta = 0$  and  $H_1 : \theta > 0$  with  $\theta$  being a scalar that defines the signal amplitude; i.e., the  $H_1$  is composite because of lack of knowledge on the signal amplitude [9–12, 14–16, 95, 108, 128, 129]; and
- The least favorable distribution approach, in which the system designer solves the composite hypothesis issue by choosing the least favorable distribution within  $H_1$ , building a simple hypothesis with just the least favorable

distribution, and using the known system design techniques for the simple hypotheses case [50, 59, 73, 123]. The resulting design is conservative in the sense that a system that satisfies a given detection performance when designed under the least favorable distribution, also satisfies the detection performance under any other distribution within  $H_1$ .

To the author's knowledge, when the measurements are conditionally dependent and the hypothesis  $H_1$  is composite, few results are available and such results are related to locally optimum approaches [12, 14, 15]. No results involving conditionally dependent hypotheses and least favorable distributions are known.



## Chapter 4

### Problem Statement and Research Questions

This dissertation focuses on the analytical treatment of sensor detection systems and addresses the problem that the current mathematical models are not amenable to analytical treatment unless the system designer adopts assumptions that are generally considered unrealistic. More specifically, **this dissertation addresses the problem that the sensor measurements are generally conditionally dependent when the sensor measurement depends on the distance between sensors and emitter, and the hypothesis  $H_1$  is composite because the distribution of the emitter location is unknown.**

#### 4.1 Problem 1: Conditionally Dependent Measurements

As described in Section 3.1, many results are available when measurements are conditionally i.i.d. and the hypotheses are simple. Under these conditions, a system designer can use the available results to design optimal centralized or distributed detection systems, apply large sample methods to evaluate systems with an asymptotically large number of sensors, evaluate the effect of communication channel errors, use censoring schemes to reduce energy consumption, and consider multiple access communication channels.

However, the assumption of conditionally i.i.d. measurements is considered by many authors to be unrealistic and only justified by convenience or by the lack of information about the correlation parameters [15, 39, 46, 90, 106, 111, 124, 134, 137, 138].

Indeed, although the measurements  $\{Z_i\}_{i=1}^K$  can be considered conditionally i.i.d. under  $H_0$  in many applications [94, 106, 111, 134], when the distribution of the measurements depends on the distances between the sensors and the emitter at random locations, the measurements  $\{Z_i\}_{i=1}^K$  are generally conditionally dependent under  $H_1$ , even if the measurement noise random variables  $\{W_i\}_{i=1}^K$  are i.i.d. [11, 106].

To illustrate the conditional dependence of  $\{Z_i\}_{i=1}^K$ , recall the model of Chapter 2 in which  $Z_i = A_i + W_i$  when  $\theta = \theta_1$ . Let  $A_i = \xi(\|L_i - L_e\|)$ , where  $L_e$  is the random variable corresponding to the emitter location,  $\{L_i\}_{i=1}^K$  are independent random variables corresponding to the sensor locations, and  $\xi$  is a given function. Let  $\{L_i\}_{i=1}^K$  be independent of  $L_e$ . Let  $\{W_i\}_{i=1}^K$  be i.i.d. random variables that are independent of  $\{L_i\}_{i=1}^K$  and  $L_e$  and have the cumulative

distribution function (c.d.f.)  $F_W(w)$ . Consider  $K = 2$ ; i.e., just two sensors. In this case, the joint c.d.f. of  $Z_1$  and  $Z_2$  is given by

$$\begin{aligned}
P[Z_1 \leq z_1, Z_2 \leq z_2 | \theta = \theta_1] &= \int_{S_e} P[Z_1 \leq z_1, Z_2 \leq z_2 | L_e = l_e, \theta = \theta_1] dP_{L_e}(l_e) \\
&= \int_{S_e} \int_{S_s} \int_{S_s} P[Z_1 \leq z_1, Z_2 \leq z_2 | L_1 = l_1, L_2 = l_2, L_e = l_e, \theta = \theta_1] dP_{L_1}(l_1) dP_{L_2}(l_2) dP_{L_e}(l_e) \\
&= \int_{S_e} \left( \int_{S_s} P[\xi(\|l_1 - l_e\|) + W_1 \leq z_1 | L_1 = l_1, L_e = l_e, \theta = \theta_1] dP_{L_1}(l_1) \right. \\
&\quad \cdot \left. \int_{S_s} P[\xi(\|l_2 - l_e\|) + W_2 \leq z_2 | L_2 = l_2, L_e = l_e, \theta = \theta_1] dP_{L_2}(l_2) \right) dP_{L_e}(l_e) \\
&= \int_{S_e} \left( \int_{S_s} F_W(z_1 - \xi(\|l_1 - l_e\|)) dP_{L_1}(l_1) \int_{S_s} F_W(z_2 - \xi(\|l_2 - l_e\|)) dP_{L_2}(l_2) \right) dP_{L_e}(l_e) \\
&= \int_{S_e} E[F_W(z_1 - \xi(\|L_1 - l_e\|))] \cdot E[F_W(z_2 - \xi(\|L_2 - l_e\|))] dP_{L_e}(l_e), \tag{4.1}
\end{aligned}$$

which is different from  $E[F_W(z_1 - \xi(\|L_1 - L_e\|))] \cdot E[F_W(z_2 - \xi(\|L_2 - L_e\|))] = P[Z_1 \leq z_1 | \theta = \theta_1] \cdot P[Z_2 \leq z_2 | \theta = \theta_1]$  in general.

Furthermore, the distribution of the measurement  $Z_i$  under  $H_1$  is a mixture of the measurement noise distribution  $F_W$ , preventing the use of common tools to model conditional dependence. This can be seen from the example in the previous paragraph, in which

$$P[Z_1 \leq z_1 | \theta = \theta_1] = \int_{S_e} \int_{S_s} F_W(z_1 - \xi(\|l_1 - l_e\|)) dP_{L_1}(l_1) dP_{L_e}(l_e); \tag{4.2}$$

therefore, even if  $Z_i$  conditioned on  $\{L_i = l_i, L_e = l_e, \theta = \theta_1\}$  is Gaussian distributed, the mixture will not be Gaussian distributed. Therefore, it is not possible to capture the dependence among  $\{Z_i\}_{i=1}^K$  through the multivariate Gaussian distribution, which is a common tool to handle conditionally dependent situations as discussed in Section 3.2.

When sensors are in known deterministic locations, there are approaches to circumvent the conditional dependence problem [98, 134]; however, these approaches require more complicated sensor operating schemes. As described in Section 3.2.1.1, the authors in [98] propose that sensors initially estimate the emitter's location; and the approach proposed in [134] requires the region of interest to be partitioned in several subregions and sensors need to follow a schedule to periodically reconfigure themselves to collect measurements from each subregion at a time.

When the measurements  $\{Z_i\}_{i=1}^K$  are conditionally dependent, it is more difficult to determine optimal designs, as discussed in detail in Section 3.2, and many of the available results reached under the conditionally i.i.d. assumption may no longer hold.

## 4.2 Problem 2: Composite Hypothesis

As discussed in detail in Section 3.3, when the hypothesis  $H_1$  is composite, the analytical treatment is further complicated because a Uniformly Most Powerful (UMP) detector may not exist [73] and because the designer has to ensure that a prescribed performance is achieved for each of the uncountably many distributions within  $H_1$ .

In the detection problem being considered in this dissertation, the amplitude of the signal depends on the distance between sensors and emitter; therefore, the hypothesis  $H_1$  is composite when the distribution of the emitter location is unknown. More precisely, consider the model of Chapter 2 and assume that the distributions for sensor locations  $\{L_i\}_{i=1}^K$ , the measurement noise distribution  $F_W$ , and the amplitude function  $\xi$  are known or given. The measurement  $Z_i$  under  $H_1$  depends on the random variable  $A_i$  that has a distribution that depends on the distance between the sensor location  $L_i$  and the emitter location  $L_e$ . It is reasonable to assume that the distribution of  $L_i$  is given since the sensor deployment strategy is imposed by the system designer; however, it may not be reasonable to assume that the distribution of  $L_e$  is given. For instance, when detecting a low-level point radiation source [20, 58, 87, 98], the distribution of where the source will be released is unavailable. Since each of the possible distributions for  $L_e$  induces a distribution for the measurements  $\{Z_i\}_{i=1}^K$ , the hypothesis  $H_1$  is composite.

As described in Section 3.3, three general approaches were considered in the literature to design a sensor detection system: the generalized likelihood ratio test (GLRT), locally optimum approaches, and least favorable distributions.

The GLRT approach is not applicable when sensors are not equipped with procedures to estimate the emitter location  $L_e$ .<sup>1</sup> It is also difficult to estimate the distribution for  $L_e$  when the presence of the signal emitter is rare, as is the case in many of the applications described in Section 1.1.

The locally optimum approach is not suitable for the setting being considered here because it requires that the hypothesis  $H_1$  be composite due to one and only one scalar parameter  $\theta$  that defines the signal amplitude, which means that all sensors would be sensing the emitter with the same amplitude. In the setting considered in this dissertation, a minimum known level for the signal amplitude is assumed and  $H_1$  is composite because of the lack of knowledge about the emitter location.

The approach of a least favorable distribution is appealing because it reduces  $H_1$  to a simple hypothesis, and if the design under this distribution meets the prescribed detection performance requirement, then it also satisfies the detection performance under any other distribution within  $H_1$ . There are, however, three issues associated with this approach:

- the theory of least favorable distributions cannot be readily applied because, as discussed in Section 3.3.2, most of the existing theory relies on the assumption of conditionally independent or conditionally i.i.d. measurements and most of the results discuss least favorable distributions for the measurements and not for a particular unknown parameter such as the emitter location  $L_e$ ;
- the approach of least favorable distributions may lead to overly conservative designs.

A system designer may be tempted to assume an arbitrary distribution for the emitter location  $L_e$ ; for instance, given the lack of knowledge, it is natural to assume that  $L_e$  is uniformly distributed within the region of interest. Such

---

<sup>1</sup>Note that in order to estimate the location of the emitter, sensors must also know their own location, which may require additional equipment in the sensors.

an assumption would make the hypothesis  $H_1$  simple; however, the resulting system could fail to meet the prescribed detection performance under the actual distribution for  $L_e$ .

### 4.3 Research Questions

Motivated by the difficulties imposed by conditionally dependent measurements, the following question is considered in this dissertation: **Are there situations in which the system designer is justified in assuming that the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d.?** If such situations exist, then a system designer is able to benefit from much of the research available for conditionally i.i.d. measurements. This question is answered in Chapter 6.

In order to address the composite hypothesis issue, the approach of least favorable distributions is applied to the distribution of the emitter location  $L_e$  and the following question is considered: **What are the conditions for finding a least favorable distribution for the emitter location  $L_e$ ?** Being able to find such a distribution will avoid the problems associated with the composite hypothesis  $H_1$ . This question is answered in Chapter 7, where it is shown that the distribution that places the emitter on the boundary of certain regions of interest is least favorable for several systems of interest.

Given the concern that least favorable distributions may lead to too conservative designs, the following question is considered in this dissertation: **If a system designer builds a sensor detection system using a least favorable emitter location distribution for a system, how conservative is such a design?** This question is answered in Chapter 8.

Given the form of the least favorable distribution for the emitter location identified in Chapter 7, which places the emitter on the boundary of certain regions of interest, a system designer may wonder whether to deploy sensors in a larger region in order to surround the boundary of the region of interest with more sensors. Given this option, the following question is considered: **Would the deployment of sensors in a region larger than the region of interest improve the detection performance under the least favorable distribution?** The answer to this question is provided in Chapter 9.

## Chapter 5

### Main Assumptions and Considerations

Most of the propositions in this dissertation focus on the sensor detection system model presented in Chapter 2 and rely on the following additional assumptions.

**Assumption I:** The measurement noise random variables  $\{W_i\}_{i=1}^K$  are i.i.d. random variables with a common c.d.f.  $F_W$ . The random variables  $\{W_i\}_{i=1}^K$  and  $\theta$  are independent. These assumptions can be considered a good approximation when detecting point radiation sources [20, 106] or transmissions in a cooperative spectrum-sensing system [120, 139], and were also adopted in [14, 90, 108, 118].

**Assumption II:** The region of interest  $S_e$  is a subset of  $\mathbb{R}^\delta$ , for some integer  $\delta \geq 1$ , and contains at most a single emitter, which is the worst case for the detection problem for many scenarios of interest, as shown in Appendix A.

**Assumption III:** The emitter location  $L_e$  is a random variable within the region of interest  $S_e$ . No prior knowledge of  $L_e$  or its distribution is assumed. The random variables  $L_e$ ,  $\{W_i\}_{i=1}^K$ , and  $\theta$  are independent.

**Assumption IV:** The sensor locations  $\{L_i\}_{i=1}^K$  are independent random variables within the deployment region  $S_s$ , and  $\{L_i\}_{i=1}^K$ ,  $L_e$ ,  $\{W_i\}_{i=1}^K$ , and  $\theta$  are independent. The case in which  $\{L_i\}_{i=1}^K$  are random variables is of interest in certain scenarios and was considered by many previous authors, as discussed in Appendix B.

**Assumption V:** For each  $i$ ,  $A_i$  and  $\{W_i\}_{i=1}^K$  are independent, and, for any  $i \neq j$ ,  $A_i$  and  $L_j$  are independent. The realizations of  $\{A_i\}_{i=1}^K$  are nonnegative and, for any  $i$  and any measurable set  $\mathcal{A}$ ,

$$P[A_i \in \mathcal{A} | \theta, L_i, L_e] := v(\mathcal{A} | \theta \cdot \xi(\mathbf{dist}(L_i, L_e))), \quad (5.1)$$

where  $v(\cdot|x)$  is a deterministic function that takes a measurable set and a real value as input, defines a measurable function of  $x$  for any measurable set, and defines a probability measure for any value  $x \in \mathbb{R}$ ;  $\mathbf{dist}(L_i, L_e)$  is a function that represents the distance between  $L_i$  and  $L_e$ ; and  $\xi$  is a deterministic known *amplitude function* that models the decay in signal amplitude as the distance between sensor and emitter increases. In most of

the developments in this dissertation,  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$ , where  $\|\cdot\|$  represents the Euclidean norm; i.e., letting  $l \in \mathbb{R}^\delta$  be a column-vector and  $l'$  be its transpose,  $\|l\| := \sqrt{l'l}$ . In other developments, when considering that  $S_e$  is a sphere in  $\mathbb{R}^3$ ,  $\mathbf{dist}(L_i, L_e)$  is defined as the minimum length of all curves that are contained in the sphere and connect  $L_i$  and  $L_e$ .

These assumptions allow the possibility that  $A_i = \theta \cdot \xi(\mathbf{dist}(L_i, L_e))$ , or that  $A_i$  has a conditional Poisson distribution with parameter  $\theta \cdot \xi(\mathbf{dist}(L_i, L_e))$ .<sup>1</sup>

Definition (5.1) implies that, for any  $i$  and  $j$ , whenever  $\xi(\mathbf{dist}(L_i, L_e)) = \xi(\mathbf{dist}(L_j, L_e))$ ,  $A_i$  and  $A_j$  are conditionally identically distributed.

**Assumption VI:** For any  $\mathbf{l} := (l_1, \dots, l_K)$  such that  $l_i \in S_s$  for all  $i$  and any  $l_e \in S_e$ ,<sup>2</sup>

$$P\left[\bigcap_{i=1}^K A_i \leq a_i \mid \theta = \theta_1, \mathbf{L} = \mathbf{l}, L_e = l_e\right] := \prod_{i=1}^K P[A_i \leq a_i \mid \theta = \theta_1, L_i = l_i, L_e = l_e]; \quad (5.2)$$

It is important to observe that, even with Assumption VI,  $\mathbf{A} := (A_1, \dots, A_K)$  is a vector of dependent random variables in general. Note further that Assumption VI is readily justified when  $A_i = \theta \cdot \xi(\mathbf{dist}(L_i, L_e))$ .<sup>3</sup> Even when  $A_i$  conditioned on  $\theta = \theta_1$ ,  $L_i = l_i$ , and  $L_e = l_e$  is not a constant random variable, Assumption VI is justified in certain detection problems, such as detection of low-level point radiation sources [20, 106].

**Assumption VII:** The amplitude function  $\xi$  is a bounded, nonnegative, nonincreasing, and right-continuous function.

**Assumption VIII:** Let  $\mu$  be either Lebesgue measure on  $\mathbb{R}$  or counting measure on  $\mathbb{Z}$ . The random variable  $W$  has probability measure that is absolutely continuous with respect to  $\mu$ . Let  $f_W(w)$  represent the density function of  $W$  with respect to  $\mu$ . The conditional distribution  $P_{Z_i|\theta=\theta_1}$  is absolutely continuous with respect to  $P_{Z_i|\theta=0}$ .

From Assumption VIII, it follows that  $Z_i$  also has a density function  $f_{Z_i|\theta=0}(z)$  with respect to  $\mu$  when conditioned on  $\theta = 0$  and a density function  $f_{Z_i|\theta=\theta_1, L_e \in S'_e}(z)$  when conditioned on  $\{\theta = \theta_1, L_e \in S'_e\}$  for any  $S'_e$  in which  $P[L_e \in S'_e] > 0$  or any  $S'_e = \{l_e\}$  for  $l_e \in S_e$ , as shown in Appendix E.

Both centralized and distributed detection systems will be considered.

Centralized detection systems are of interest because they can approximate many envisioned sensor detection systems. Although the sensing device or the communication subsystem may include a quantization phase in which the measurement is represented by a finite number of bits, the approximation can be considered good when the number of bits is large enough. Many of the wireless network nodes and protocols envisioned to build a network of sensors offer

<sup>1</sup>More precisely, this means that  $P[A_i = n \mid \theta, L_i, L_e] = (\theta \xi(\mathbf{dist}(L_i, L_e)))^n \exp\{-\theta \xi(\mathbf{dist}(L_i, L_e))\}/n!$ .

<sup>2</sup>Note that the definition of (5.2) ensures that the function defined by  $P[\bigcap_{i=1}^K \{A_i \leq a_i\} \mid \theta = \theta_1, \mathbf{L} = \mathbf{l}, L_e = l_e]$  is a joint distribution for any  $\theta_1 > 0$ , any  $\mathbf{l}$ , and any  $l_e$ .

<sup>3</sup>To see this, observe that, when  $A_i = \theta \cdot \xi(\mathbf{dist}(L_i, L_e))$ ,  $A_i$  given  $\theta = \theta_1$ ,  $L_i = l_i$ , and  $L_e = l_e$  is a constant random variable and constant random variables are independent.

mechanisms for the reliable transfer of data packets offering several bytes to represent the measurement [68, 98, 104, 110, 140].

The many bytes offered by wireless networks and protocols may suggest that the tight quantization constraints normally considered in a distributed detection system are artificial; however, there are still reasons to reduce the amount of information generated by sensors. One reason is that the sensor detection system may share the communication subsystem with other systems that occupy most of the available capacity; for instance, the authors in [64] describe the deployment of a sensor network in which the various bits available in a data packet are allocated for various different information systems and a single bit is reserved for the detection system. Another reason for considering tight quantization constraints is that it may be more suitable for multiple access channels; for instance, as described in Section 3.1.5, some authors have considered a communication subsystem in which sensors produce a binary output to decide whether to transmit a common waveform to the fusion center, which receives the sum of the received signals. The use of a binary output enables the expected value of the received signal to have an amplitude proportional to the number of sensors that transmit the waveform [78]. Lastly, tight quantization constraints are still being considered by recent papers targeting specific applications [105–107].

## Chapter 6

### Scenarios for Conditionally Independent and Conditionally I.I.D. Measurements

Motivated by the difficulty of designing sensor detection systems under conditionally dependent measurements, this section presents several propositions that identify conditions under which a system designer is justified in assuming that the measurements  $\mathbf{Z}$  are either conditionally independent or conditionally i.i.d. More specifically, three settings are considered: when the distribution of the emitter location  $L_e$  is supported within a suitable  $S_e^* \subset S_e$ ; when the amplitude function  $\xi(\cdot)$  vanishes; and when  $S_e$  is a sphere. The first setting may seem artificial at first; however, as shown in subsequent chapters, it will be useful for the design of a sensor detection system.

#### 6.1 Emitter Location Distributed on a Specific Subset of the Region of Interest

The propositions in this subsection show that, if the emitter location is distributed on a specific subset  $S_e^*$  of  $S_e$ , then the conditional dependence in the measurements disappears.

A simple scenario in which the measurements  $\{Z_i\}_{i=1}^K$  become conditionally independent given  $\theta$  is the one in which the sensor positions  $\{L_i\}_{i=1}^K$  are deterministic and  $S_e^*$  contains a single element; i.e., when  $L_e$  is a constant random variable. This scenario has been considered by other researchers [118] and the conclusion that the measurements become conditionally independent given  $\theta$  can be obtained as a special case of Lemma C.1 in Appendix C.

Even when the sensor positions  $\{L_i\}_{i=1}^K$  are i.i.d. random variables, it follows from (4.1) in Chapter 4 that the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  when  $S_e^*$  has a single element.

However, as shown in Chapter 4, the measurements  $\{Z_i\}_{i=1}^K$  will not be conditionally i.i.d. given  $\theta$  in general when the range of the random variable  $L_e$  has more than one possible emitter location; i.e., when  $S_e^*$  has more than one element.

Nevertheless, there are sufficient conditions under which  $\{Z_i\}_{i=1}^K$  become conditionally i.i.d. given  $\theta$  when  $S_e^*$  has more than one element. As shown in the next lemma, an important sufficient condition is to have the conditional



distribution given  $L_e = l_e$  of the parameter  $\xi(\mathbf{dist}(L_i, L_e))$  that defines the distribution of  $A_i$  be invariant to changes in  $l_e \in S_e^*$ .

**Lemma 6.1.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, and VI of Chapter 5. Let  $X_i := \xi(\mathbf{dist}(L_i, L_e))$ , where  $\xi(\mathbf{dist}(L_i, L_e))$  defines the distribution of  $A_i$  as per Assumption V. If*

- *the distribution of  $L_e$  satisfies  $P[L_e \in S_e^*] = 1$  for some  $S_e^* \subseteq S_e$ ;*
- *the sensor locations  $\{L_i\}_{i=1}^K$  are independent random variables;*
- *the conditional distribution of  $X_i$  given  $L_e = l_e$  is invariant to changes in  $l_e \in S_e^*$  for any  $x$ ; i.e.,*

$$\forall i, \forall l_{e_1}, l_{e_2} \in S_e^*, \forall x, P[X_i \leq x | L_e = l_{e_1}] = P[X_i \leq x | L_e = l_{e_2}] \quad (6.1)$$

*then  $\{Z_i\}_{i=1}^K$  are conditionally independent given  $\theta$  and are independent of  $L_e$ . Additionally, if  $\{L_i\}_{i=1}^K$  are i.i.d., then  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$ .*

*Proof:* To show that  $\{Z_i\}_{i=1}^K$  and  $L_e$  are independent, it is first shown that  $\{Z_i\}_{i=1}^K$  and  $L_e$  are conditionally independent given either  $\theta = 0$  or  $\theta = \theta_1$ . Once this is proven, pick any  $\{z_i\}_{i=1}^K$  and any measurable set  $\mathcal{B} \subset \mathbb{R}^\delta$ , and write

$$\begin{aligned} P\left[\bigcap_{i=1}^K \{Z_i \leq z_i\}, L_e \in \mathcal{B}\right] &= \sum_{j \in \{0, \theta_1\}} P[\theta = j] P\left[\bigcap_{i=1}^K \{Z_i \leq z_i\}, L_e \in \mathcal{B} \mid \theta = j\right] \\ &= \sum_{j \in \{0, \theta_1\}} P[\theta = j] P\left[\bigcap_{i=1}^K \{Z_i \leq z_i\} \mid \theta = j\right] P[L_e \in \mathcal{B} \mid \theta = j] \\ &= P[L_e \in \mathcal{B}] \sum_{j \in \{0, \theta_1\}} P[\theta = j] P\left[\bigcap_{i=1}^K \{Z_i \leq z_i\} \mid \theta = j\right] = P[L_e \in \mathcal{B}] \cdot P\left[\bigcap_{i=1}^K \{Z_i \leq z_i\}\right], \end{aligned} \quad (6.2)$$

where (6.2) follows because  $L_e$  is independent of  $\theta$ .

The conclusion that  $\{Z_i\}_{i=1}^K$  and  $L_e$  are conditionally independent given  $\theta = 0$  follows from the additive model (2.1) and Assumptions I and III.

For  $\theta = \theta_1$ ,  $\{Z_i\}_{i=1}^K$  and  $L_e$  will be conditionally independent given  $\theta = \theta_1$  if and only if  $\{A_i\}_{i=1}^K$  and  $L_e$  are conditionally independent given  $\theta = \theta_1$ . For the same reasons,  $\{Z_i\}_{i=1}^K$  will be conditionally i.i.d. given  $\theta = \theta_1$  if and only if  $\{A_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta = \theta_1$ .

The conclusion that  $\{A_i\}_{i=1}^K$  and  $L_e$  are conditionally independent given  $\theta = \theta_1$  follow once two statements are proven: (1)  $\{A_i\}_{i=1}^K$  are conditionally independent given  $\{\theta = \theta_1, L_e = l_e\}$ ; and (2) for any  $i$ ,  $P[A_i \leq a_i \mid \theta = \theta_1, L_e = l_e] := v_i(a_i)$  is invariant to changes in  $l_e \in S_e^*$ . If these two conditions are satisfied, then, as shown in details in Lemma C.1 in the appendix, for any  $\{a_i\}_{i=1}^K$  and any measurable set  $\mathcal{B} \subset \mathbb{R}^\delta$ ,  $v_i(a_i) = P[A_i \leq a_i \mid \theta = \theta_1]$ , and

$$P\left[\bigcap_{i=1}^K \{A_i \leq a_i\}, L_e \in \mathcal{B} \mid \theta = \theta_1\right] = \int_{\mathcal{B} \cap S_e^*} P\left[\bigcap_{i=1}^K A_i \leq a_i \mid \theta = \theta_1, L_e = l_e\right] dP_{L_e}(l_e) \quad (6.3)$$

$$= \int_{\mathcal{B} \cap S_e^*} \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1, L_e = l_e] dP_{L_e}(l_e) \quad (6.4)$$

$$= \int_{\mathcal{B} \cap S_e^*} \prod_{i=1}^K v_i(a_i) dP_{L_e}(l_e) = \prod_{i=1}^K v_i(a_i) P[L_e \in \mathcal{B} \cap S_e^*] \\ = \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1] \cdot P[L_e \in \mathcal{B} | \theta = \theta_1] \quad (6.5)$$

and  $\{A_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta = \theta_1$  if  $\{L_i\}_{i=1}^K$  are identically distributed.

The condition that  $\{A_i\}_{i=1}^K$  be conditionally independent given  $\{\theta = \theta_1, L_e = l_e\}$  follows from Assumption VI and because  $\{L_i\}_{i=1}^K$  are conditionally independent. For a detailed derivation, refer to Lemma C.2 in the appendix.

The condition that  $P[A_i \leq a_i | \theta = \theta_1, L_e = l_e]$  is invariant to changes in  $l_e \in S_e^*$  for any  $a_i$  follows from Assumption V in that the distribution of  $A_i$  depends on  $L_e$  through the parameter  $X_i$ ; and, as shown in details in Lemma C.3, the condition follows if  $P[X_i \leq x | L_e = l_e]$  is invariant to changes in  $l_e \in S_e^*$  for any  $x$ , which is assumed in this lemma. ■

Lemma 6.1 is now used to establish various sufficient conditions under which  $\{Z_i\}_{i=1}^K$  become conditionally i.i.d. given  $\theta$  when  $S_e^*$  has more than one element.

**Proposition 6.2.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, and VI of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V, and let  $B_r(0)$  denote the closed ball centered at the origin and having radius  $r$ . Let  $\partial B_r(0)$  denote the boundary of  $B_r(0)$ . If*

- *the region of interest  $S_e = B_{r_M}(0)$  for some  $r_M > 0$ ;*
- *the distribution of  $L_e$  satisfies  $P[L_e \in S_e^*] = 1$  for  $S_e^* = \partial B_r(0)$  for some  $r \in [0, r_M]$ ;*
- *the sensor locations  $\{L_i\}_{i=1}^K$  are independent random variables distributed in  $S_s = B_{c \cdot r_M}(0)$  for some  $c > 0$ ;*
- *For any  $i$ ,  $L_i$  has a distribution invariant to orthogonal transformations (e.g. rotations); i.e., for any measurable set  $\mathcal{B} \subset \mathbb{R}^\delta$  and for any orthogonal  $\delta$ -by- $\delta$  matrix  $H$ ,  $P[L_i \in \mathcal{B}] = P[L_i \in H\mathcal{B}]$ ;*

*then  $\{Z_i\}_{i=1}^K$  are conditionally independent given  $\theta$  and are independent of  $L_e$ . Additionally, if  $\{L_i\}_{i=1}^K$  are i.i.d., then  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$ .*

*Proof:* To prove the conclusions of this propositions, it is enough to prove that the conditions of this proposition imply the conditions of Lemma 6.1. The conditions on  $L_e$  and on  $\{L_i\}_{i=1}^K$  of Lemma 6.1 are also conditions of this proposition; thus, it remains to prove that  $P[X_i \leq x | L_e = l_e]$  is invariant to changes in  $l_e \in S_e^*$  for any  $x$ ; i.e., condition (6.1) for  $X_i = \xi(\|L_i - L_e\|)$ , for the specified distribution of  $L_i$  and the specified sets  $S_e$ ,  $S_e^*$ , and  $S_s$ .

Since  $X_i$  is a function of  $L_e$  through  $D_i := \|L_i - L_e\|$ , (6.1) for  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  follows if  $P[D_i \leq d | L_e = l_e]$  is invariant to changes in  $l_e \in S_e^*$  for any  $d$ ; i.e.,

$$\forall i, \forall l_{e_1}, l_{e_2} \in S_e^*, \forall d, P[D_i \leq d | L_e = l_{e_1}] = P[D_i \leq d | L_e = l_{e_2}], \quad (6.6)$$

as shown in details in Lemma C.4 in the appendix.

In order to prove (6.6), observe that  $P[D_i \leq d | L_e = l_e] = P[\|L_i - l_e\| \leq d] = P[L_i \in B_d(l_e) \cap S_s]$ , and (6.6) follows if

$$\forall i, \forall l_{e_1}, l_{e_2} \in S_e^*, \forall d, P[L_i \in B_d(l_{e_1}) \cap S_s] = P[L_i \in B_d(l_{e_2}) \cap S_s]. \quad (6.7)$$

To see that (6.7) follows, observe that for any  $l_{e_1}, l_{e_2} \in S_e^*$ , there is a rotation that transforms  $l_{e_1}$  into  $l_{e_2}$ , and since both the set  $S_s$  and the distribution of  $L_i$  are invariant to orthogonal transformations, which include rotations, (6.7) follows. This argument is made precise in Lemma C.5 in the appendix. To use Lemma C.5, let  $\mathcal{H}^*$  be the set of all orthogonal transformations. The first condition of Lemma C.5 is satisfied because, for any  $l_{e_1}, l_{e_2} \in S_e^*$ , there exists an orthogonal matrix  $H_{1,2}$  for which  $H_{1,2}l_{e_1} = l_{e_2}$ . To see this, for either  $j = 1$  or  $j = 2$ , let  $H_j := [l_{e_j}/\|l_{e_j}\| \mid v_{j,2} \mid \dots \mid v_{j,\delta}]$ , where  $\{v_{j,k}\}_{k=2}^\delta$  are vectors such that  $l_{e_j}/\|l_{e_j}\|$  and  $\{v_{j,k}\}_{k=2}^\delta$  form a basis for  $\mathbb{R}^\delta$ , and let  $H_{1,2} = H_2 H_1^t$ . Observe that  $H_2 H_1^t$  is also orthogonal since both  $H_1$  and  $H_2$  are orthogonal. The invariance to orthogonal transformations of  $S_e^*$ ,  $S_s$ , and  $S_e$  is proven in Lemma C.6 in the appendix, which means that the second condition of Lemma C.5 is satisfied. The third condition of Lemma C.5 is satisfied by assumption. ■

Observe that the uniform distribution of  $L_i$  is a distribution invariant to orthogonal transformations. To see this, let  $H$  be any orthogonal  $\delta$ -by- $\delta$  matrix, pick any measurable  $\mathcal{B}$ , let  $m$  be Lebesgue measure on  $\mathbb{R}^\delta$ , and write

$$P[L_i \in \mathcal{B}] = \frac{m(\mathcal{B} \cap S_s)}{m(S_s)} = \frac{m(H(\mathcal{B} \cap S_s))}{m(S_s)} \quad (6.8)$$

$$= \frac{m(H\mathcal{B} \cap HS_s)}{m(S_s)} = \frac{m(H\mathcal{B} \cap S_s)}{m(S_s)} = P[L_i \in H\mathcal{B}] \quad (6.9)$$

where (6.8) follows because, for any linear transformation  $T$ ,  $m(T(\mathcal{B})) = |\det(T)|m(\mathcal{B})$  [47, p. 74] and  $|\det(H)| = 1$  [51, p. 208], the first equality of (6.9) follows because any matrix  $H \in \mathcal{H}$  is invertible, which means that  $H$  forms a bijective transformation and, therefore,  $H(S_1 \cap S_2) = H(S_1) \cap H(S_2)$ , and the second equality of (6.9) follows because  $HS_s = S_s$ .

Considering the special case in which the region of interest  $S_e$  is in the plane, the next proposition provides sufficient conditions for  $\{Z_i\}_{i=1}^K$  to be conditionally i.i.d. given  $\theta$  in non-circular regions of interest.

**Proposition 6.3.** *Adopt the additive model of (2.1), Assumptions I, II, III, IV, V, and VI of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V and let  $S_e \subset \mathbb{R}^2$ . If*

- *the region of interest  $S_e$  is a regular convex polygon with  $n$  vertices<sup>1</sup> and with a circumscribing circle  $\partial B_{r_M}(0)$  centered at the origin and with radius  $r_M$ ;*

<sup>1</sup>A polygon with  $n$  vertices is defined as the union of the set of  $n$  points  $\{v_i\}_{i=0}^{n-1}$  in  $\mathbb{R}^2$ , which are called the vertices of the polygon, the set of line segments  $\mathcal{V} := \{v_0 v_1, v_1 v_2, \dots, v_{n-1} v_0\}$  between these points, which are called the sides of the polygon, and the interior of the set formed by such line segments. Note that the set of line segments  $\mathcal{V}$  decomposes  $\mathbb{R}^2$  in two sets: one set with infinite area, and another set with finite area, which is called the interior set. A convex polygon is a polygon whose interior set is convex. A regular convex polygon is a polygon that is equiangular and equilateral; i.e., the angle formed by the two sides that contain a vertex is equal to a constant for all vertices, and all sides have the same length [34]. Examples of regular convex polygons are the equilateral triangle and the square.

- the distribution of  $L_e$  satisfies  $P[L_e \in S_e^*] = 1$  for  $S_e^* = c_1(S_e \cap \partial B_{r_M}(0))$  for any  $c_1 \in [0, 1]$ ; i.e.,  $S_e^*$  is the set formed by all the vertices of  $S_e$  multiplied by a scaling factor  $c_1$ ;
- the deployment region satisfies  $S_s = B_r(0)$  for any  $r > 0$  or  $S_s = c_2 \cdot S_e$ , for any  $c_2 > 0$ ; i.e.,  $S_s$  is either a disk or  $S_s$  is a scaled version of  $S_e$ ;
- $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution on  $S_s$ ;

then  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  and are independent of  $L_e$ .

*Proof:* As in the proof of Proposition 6.2, it is enough to prove that the conditions of this proposition imply the conditions of Lemma 6.1. The conditions on  $L_e$  and on  $\{L_i\}_{i=1}^K$  of Lemma 6.1 are also conditions of this proposition; thus, it remains to prove that  $P[X_i \leq x | L_e = l_e]$  is invariant to changes in  $l_e \in S_e^*$  for any  $x$ ; i.e., condition (6.1) for  $X_i = \xi(\|L_i - L_e\|)$ , for the specified distribution of  $L_i$  and the specified sets  $S_e$ ,  $S_e^*$ , and  $S_s$ . Also as in the proof of Proposition 6.2, since  $X_i$  is a function of  $L_e$  through  $D_i := \|L_i - L_e\|$ , the conclusions are reached once the condition that  $P[D_i \leq d | L_e = l_e]$  be invariant to changes in  $l_e \in S_e^*$  for any  $d$ ; i.e., (6.6) for  $D_i = \|L_i - L_e\|$ , is proven.

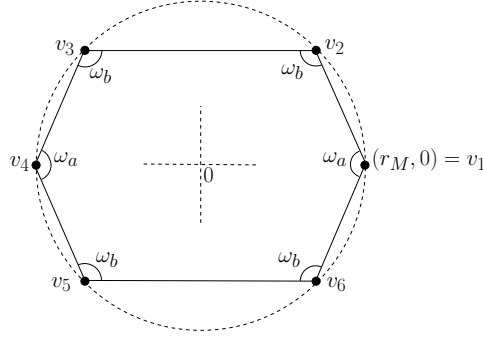
Recalling that  $D_i$  conditioned on  $L_e = l_e$  equals  $\|L_i - l_e\|$  and  $P[\|L_i - l_e\| \leq d] = P[L_i \in B_d(l_e) \cap S_s]$ , (6.6) follows because the distribution of  $L_i$  and the set  $S_s$  are invariant to orthogonal transformations that preserve  $S_e^*$ . More precisely, (6.6) for the  $S_e$ ,  $S_s$ , and  $S_e^*$  of this proposition follows from Lemma C.5 in the appendix. To use Lemma C.5, let  $n$  be the number of vertices of the polygon that define  $S_e$  and let  $\mathcal{H}^* := \{H_j : j \in \{0, 1, \dots, n-1\}\}$ , with  $H_j$  given by

$$H_j := \begin{bmatrix} \cos(2\pi j/n) & \sin(2\pi j/n) \\ -\sin(2\pi j/n) & \cos(2\pi j/n) \end{bmatrix}. \quad (6.10)$$

The first and second conditions of Lemma C.5 are satisfied because, as shown in Lemma C.7 in the appendix, for any  $l_{e_1}, l_{e_2} \in S_e^*$ , there exists an orthogonal matrix  $H_{1,2} \in \mathcal{H}^*$  for which  $H_{1,2}l_{e_1} = l_{e_2}$ ; and because  $S_e$ ,  $S_e^*$ , and  $S_s$  are invariant to the transformation of any  $H_j$ . The third condition of Lemma C.5 is satisfied because the uniform distribution of  $L_i$  is invariant to any orthogonal transformation, as shown by relations (6.8) and (6.9), and each  $H_j$  is an orthogonal transformation. ■

The conditions of Propositions 6.2 or 6.3 that the emitter be located on a circle or at the vertices of a regular polygon seem difficult to satisfy; however, as will become clear in Chapter 7, there are conditions under which having  $L_e$  be in such subsets with probability 1 is the least favorable distribution for a design.

Before further conditions that cause the measurements  $\{Z_i\}_{i=1}^K$  become conditionally i.i.d. are explored, an example is presented to show that the conclusions of Proposition 6.3 do not hold for arbitrary polygonal regions of interest and to illustrate an important concept behind Propositions 6.2 and 6.3.



**Figure 6.1.** Non-regular convex hexagonal region of interest  $S_e$  in which the distribution for the emitter location that places the emitter in each of the vertices of the polygon with probability  $1/6$  does not cause the measurements  $\{Z_i\}_{i=1}^K$  to become conditionally i.i.d.

### 6.1.1 Example 6.1:

This example illustrates that placing the emitter location at the vertices of a non-regular convex polygon with probability one does not always cause the measurements to become conditionally i.i.d. This example also highlights an important concept behind Propositions 6.2 and 6.3.

Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let  $S_e$  be the non-regular convex hexagon whose vertices are at

$$\begin{aligned} v_1 &:= (r_M, 0), & v_2 &:= (r_M \cos(\pi/4), r_M \sin(\pi/4)), & v_3 &:= (-r_M \cos(\pi/4), r_M \sin(\pi/4)), \\ v_4 &:= (-r_M, 0), & v_5 &:= (-r_M \cos(\pi/4), -r_M \sin(\pi/4)), & v_6 &:= (r_M \cos(\pi/4), -r_M \sin(\pi/4)) \end{aligned} \quad (6.11)$$

for any  $r_M > 0$ . It is possible to see that these vertices are on the circle  $\partial B_{r_M}(0)$ . Let  $\omega_a$  be the internal angle of vertices  $v_1$  and  $v_4$  and  $\omega_b$  be the internal angle of the remaining vertices, as illustrated in Figure 6.1. Let  $S_e^* = S_e \cap \partial B_{r_M}(0)$ ; i.e.,  $S_e^* = \{v_1, \dots, v_6\}$ , and assume that  $P[L_e = v_j] = 1/6$  for  $j = 1, \dots, 6$ ; i.e., the distribution for the emitter location places the emitter at each of the vertices of the polygon with equal probability. Let  $A_i = \theta_1 \xi(\|L_i - L_e\|)$  with  $\xi(r) = \mathbf{1}\{r < r_0\}$ , where  $r_0 \in (0, r_M \sin(\pi/4))$  and  $\theta_1 = 1$ . Let  $F_W(w)$  be increasing. Assume that  $S_s = S_e$ , consider two sensors, and assume that the sensor locations  $L_1$  and  $L_2$  are uniformly distributed in  $S_s$ .

With the above assumptions, the only condition of Proposition 6.3 that is not satisfied is the regularity of the polygonal region of interest. It is now shown that excluding this condition is enough to cause the measurements  $Z_1$  and  $Z_2$  to be conditionally dependent.

As shown in equation (4.1) of Section 4.1,

$$P[Z_1 \leq z_1, Z_2 \leq z_2 | \theta = 1] = \sum_{j=1}^6 E[F_W(z_1 - \xi(\|L_1 - v_j\|))] E[F_W(z_2 - \xi(\|L_2 - v_j\|))] P[L_e = v_j]. \quad (6.12)$$

Since  $\xi(r) = \mathbf{1}\{r < r_0\}$ , for either  $i = 1$  or  $i = 2$ ,

$$E[F_W(z_i - \xi(\|L_i - v_j\|))] = \int_{S_s \cap B_{r_0}^\circ(v_j)} F_W(z_i - 1) dP_{L_i} + \int_{S_s \cap B_{r_0}^\circ(v_j)^c} F_W(z_i) dP_{L_i}$$

$$\begin{aligned}
&= P[L_i \in S_s \cap B_{r_0}^\circ(v_j)]F_W(z_i - 1) + (1 - P[L_i \in S_s \cap B_{r_0}^\circ(v_j)])F_W(z_i) \\
&= \frac{m(S_s \cap B_{r_0}^\circ(v_j))}{m(S_s)} \cdot [F_W(z_i - 1) - F_W(z_i)] + F_W(z_i),
\end{aligned} \tag{6.13}$$

where  $B_r^\circ(l)$  represents the open ball of radius  $r$  and centered at  $l$ ;  $m(\mathcal{A})$  gives the area of  $\mathcal{A} \subset \mathbb{R}^2$ ; and (6.13) follows from the uniform distribution assumption for  $L_i$ .

From (6.13), it is possible to see that  $E[F_W(z_i - \xi(\|L_i - v_j\|))]$  depends on  $v_j$  through the area of  $S_s \cap B_{r_0}^\circ(v_j)$ ; i.e.,  $m(S_s \cap B_{r_0}^\circ(v_j))$ . Since  $m(S_s \cap B_{r_0}^\circ(v_j))$  is the same for the vertices  $v_2, v_3, v_5$ , and  $v_6$ , and since  $m(S_s \cap B_{r_0}^\circ(v_1)) = m(S_s \cap B_{r_0}^\circ(v_4))$ , equality (6.12) can be rewritten as

$$\begin{aligned}
P[Z_1 \leq z_1, Z_2 \leq z_2 | \theta = 1] &= \frac{1}{3} \left[ \prod_{i=1}^2 \left( \frac{m(S_s \cap B_{r_0}^\circ(v_1))}{m(S_s)} \cdot [F_W(z_i - 1) - F_W(z_i)] + F_W(z_i) \right) \right] \\
&\quad + \frac{2}{3} \left[ \prod_{i=1}^2 \left( \frac{m(S_s \cap B_{r_0}^\circ(v_2))}{m(S_s)} \cdot [F_W(z_i - 1) - F_W(z_i)] + F_W(z_i) \right) \right].
\end{aligned} \tag{6.14}$$

Using similar steps, it is possible to reach that

$$\begin{aligned}
&P[Z_1 \leq z_1 | \theta = 1] \cdot P[Z_2 \leq z_2 | \theta = 1] \\
&= \prod_{i=1}^2 \left[ \frac{1}{3} \left( \frac{m(S_s \cap B_{r_0}^\circ(v_1))}{m(S_s)} \cdot [F_W(z_i - 1) - F_W(z_i)] + F_W(z_i) \right) \right. \\
&\quad \left. + \frac{2}{3} \left( \frac{m(S_s \cap B_{r_0}^\circ(v_2))}{m(S_s)} \cdot [F_W(z_i - 1) - F_W(z_i)] + F_W(z_i) \right) \right].
\end{aligned} \tag{6.15}$$

After expanding and grouping the terms in (6.14) and (6.15), one reaches that

$$\begin{aligned}
&P[Z_1 \leq z_1, Z_2 \leq z_2 | \theta = 1] - P[Z_1 \leq z_1 | \theta = 1] \cdot P[Z_2 \leq z_2 | \theta = 1] \\
&= 2 \left( \frac{m(S_s \cap B_{r_0}^\circ(v_1)) - m(S_s \cap B_{r_0}^\circ(v_2))}{3m(S_s)} \right)^2 \prod_{i=1}^2 [F_W(z_i - 1) - F_W(z_i)].
\end{aligned} \tag{6.16}$$

Since  $F_W(w)$  is increasing,  $Z_1$  and  $Z_2$  will be conditionally independent if and only if  $m(S_s \cap B_{r_0}^\circ(v_1)) - m(S_s \cap B_{r_0}^\circ(v_2)) = 0$ , which is not satisfied in this example because  $\omega_a \neq \omega_b$ .

Note however that, if all internal angles were equal, which is a property of regular convex polygons, or if the distribution of  $L_e$  placed the emitter only on the vertices that have the same internal angle, then  $Z_1$  and  $Z_2$  would be conditionally i.i.d. given  $\theta$ .

This example highlights an important concept behind Propositions 6.2 and 6.3. As seen in the proofs of both propositions, the conditional independence among  $\{Z_i\}_{i=1}^K$  could be established because of the invariance of  $P[A_i \leq a_i | \theta = \theta_1, L_e = l_e]$  with respect to changes in  $l_e \in S_e^*$  for any  $a_i$ . When  $A_i = \mathbf{1}\{\|L_i - L_e\| < r_0\}$ , as in this example, the required invariance can be seen with little effort because  $A_i \leq a_i \Leftrightarrow L_i \notin B_{r_0}^\circ(L_e)$  for any  $a_i \in [0, 1]$ ; and  $P[A_i \leq a_i | \theta = \theta_1, L_e = l_e] = 1 - m(S_s \cap B_{r_0}^\circ(l_e))/m(S_s)$ . Thus, when  $P[L_e \in S_e^*] = 1$ , conditionally i.i.d. given  $\theta$  follows when the area of  $S_s \cap B_{r_0}^\circ(l_e)$  remains constant as one varies  $l_e$  within  $S_e^*$ . While the area of  $S_s \cap B_{r_0}^\circ(l_e)$  remains constant as one varies  $l_e$  within the boundary of a circular region of interest or within the vertices of a region of interest that is a

regular convex polygon, it may not remain constant in several other regions of interest, such as the polygonal region of this example.

In the next subsections, further conditions that ensure that  $\{Z_i\}_{i=1}^K$  are conditionally independent or conditionally i.i.d. given  $\theta$  are explored.

## 6.2 Vanishing Amplitude Functions

Assume for this section that the system designer is required to design the detection system for a given distribution of the emitter location  $L_e$ . Even if the distribution of  $L_e$  does not satisfy the conditions of Propositions 6.2, or 6.3, there are still conditions under which the sensor measurements  $\{Z_i\}_{i=1}^K$  become conditionally i.i.d. given  $\theta$  if the amplitude function vanishes at some point; i.e., if  $\xi(r) = 0$  for  $r \geq r_0$  for some  $r_0 > 0$ .

**Proposition 6.4.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. For any  $S_e$  and any distribution for  $L_e$ , if*

- *there is a  $r_0 > 0$  such that  $\xi(r) = 0$  for  $r \geq r_0$ ;*
- *the region of deployment  $S_s$  satisfies*

$$S_s \supseteq \bigcup_{l \in S_e} B_{r_0}(l); \quad (6.17)$$

- *the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution on  $S_s$ ;*

*then  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  and are independent of  $L_e$ .*

*Proof:* To prove the conclusions of this propositions, it is enough to prove that the conditions of this proposition imply the conditions of Lemma 6.1. The condition on the distribution of  $L_e$  is satisfied with  $S_e^* = S_e$ . The condition on the distributions of  $\{L_i\}_{i=1}^K$  is also a condition of this proposition. It remains to prove that  $P[X_i \leq x | L_e = l_e]$  is invariant to changes in  $l_e \in S_e^*$  for any  $x$ ; i.e., condition (6.1) for  $X_i = \xi(\|L_i - L_e\|)$ . Since  $\xi \geq 0$ , it is enough to consider  $x \geq 0$  because, for  $x < 0$ ,  $P[\xi(\|L_i - l_e\|) \leq x] = 0$  for all  $l_e \in S_e$ .

To prove (6.1), let  $m(\mathcal{B})$  represent Lebesgue measure of the set  $\mathcal{B} \subset \mathbb{R}^\delta$ , assume without loss of generality that  $\delta = 2$ , and recall that  $X_i$  conditioned on  $L_e = l_e$  is equal to  $\xi(\|L_i - l_e\|)$ . Observe that, because  $\xi(d)$  is right-continuous and nonincreasing,

$$\xi(d) \leq x \Leftrightarrow d \geq d_x \quad (6.18)$$

for any  $x \geq 0$  and some  $d_x \leq r_0$ . Using this equivalence, write

$$\begin{aligned} \forall x \geq 0, P[\xi(\|L_i - l_e\|) \leq x] &= P[\|L_i - l_e\| \geq d_x] \\ &= P[L_i \notin B_{d_x}^\circ(l_e) \cap S_s] = 1 - \frac{m(B_{d_x}^\circ(l_e) \cap S_s)}{m(S_s)} \end{aligned} \quad (6.19)$$

$$= 1 - \frac{m(B_{d_x}^\circ(l_e))}{m(S_s)} = 1 - \frac{\pi \cdot d^2}{m(S_s)} \quad (6.20)$$

where  $B_d^\circ(l)$  represents the open ball of radius  $d$  and centered at  $l$ , (6.19) follows from the assumption that  $L_i$  is uniformly distributed in  $S_s$ , and (6.20) follows from condition (6.17) and  $d_x \leq r_0$  for all  $x$ . Since the area (or volume for  $\delta > 2$ ) of a ball centered in  $l_e$  with radius  $d$  does not depend on  $l_e$ , the distribution of  $X_i$  conditioned on  $L_e = l_e$  is invariant to changes in  $l_e \in S_e$ . ■

Proposition 6.4 is of significance because it means that, when the amplitude function  $\xi(r)$  vanishes for  $r \geq r_0$  and if sensors are deployed in a deployment region  $S_s$  enlarged by  $r_0$ ,  $\{Z_i\}_{i=1}^K$  become conditionally i.i.d. given  $\theta$  for any distribution for the emitter location.

### 6.3 Sensors Deployed on a Sphere

Assume for this section that the region of interest is a sphere. An example in which such a setting may be useful is when the sensor detection system is used to collect measurements from the surface of a spherical object, such as a planet or a biological spherical organ [40, 61, 83].

**Proposition 6.5.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, and VI of Chapter 5. Let the region of interest be  $S_e = \partial B_{r_M}(0) \subset \mathbb{R}^3$ ; i.e.,  $S_e$  is a sphere, and, in Assumption V, let  $\mathbf{dist}(L_i, L_e)$  be defined as the minimum length among all the curves in  $S_e$  that connect  $L_i$  and  $L_e$ . If the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution on  $S_s = S_e$ ; then  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  and are independent of  $L_e$ .*

*Proof:* As in the proofs of previous propositions, it is enough to prove that the conditions of this proposition imply the conditions of Lemma 6.1. The condition on the distribution of  $L_e$  is satisfied with  $S_e^* = S_e$ . The condition on the distributions of  $\{L_i\}_{i=1}^K$  is also a condition of this proposition. It remains to prove that  $P[X_i \leq x | L_e = l_e]$  is invariant to changes in  $l_e \in S_e$  for any  $x$ . From Lemma C.4 in the appendix, this condition follows from the condition that  $P[D_i \leq d | L_e = l_e]$  is invariant to changes in  $l_e \in S_e$  for any  $d$ ; i.e., (6.6) for  $D_i = \mathbf{dist}(L_i, L_e)$ .

To see that  $P[D_i \leq d | L_e = l_e]$  is invariant to changes in  $l_e \in S_e$  for all  $d$ , observe that  $D_i$  conditioned on  $L_e = l_e$  is equal to  $\mathbf{dist}(L_i, l_e)$ , let  $m(\mathcal{B})$  represent the spherical measure; i.e.,  $m(\mathcal{B})$  is the surface area of  $\mathcal{B} \cap S_s$  for a measurable set in  $\mathbb{R}^3$ , and write

$$\begin{aligned} \forall d, P[\mathbf{dist}(L_i, l_e) \leq d] &= P[L_i \in \{l \in S_s : \mathbf{dist}(l, l_e) \leq d\}] \\ &= \frac{m(\{l \in S_s : \mathbf{dist}(l, l_e) \leq d\})}{m(S_s)} \end{aligned} \quad (6.21)$$

where (6.21) follows from the assumption that  $L_i$  is uniformly distributed in  $S_s$ . Since the surface area of  $\{l \in S_s : \mathbf{dist}(l, l_e) \leq d\}$  does not change as one varies  $l_e \in S_e$  because  $S_s = S_e = \partial B_{r_M}(0)$ , the distribution of  $D_i$  conditioned on  $L_e = l_e$  is invariant to changes in  $l_e \in S_e$ . ■



Proposition 6.5 is significant because it means that, when the region of interest is a sphere and the signal amplitude decays with the length of curves in the sphere boundary,  $\{Z_i\}_{i=1}^K$  become conditionally i.i.d. given  $\theta$  for any distribution for the emitter location.

## Chapter 7

### Least and Most Favorable Distributions for Emitter Location

As mentioned in Chapter 4, one of the reasons why the optimal set of sensor and fusion functions is difficult to determine is the lack of *a priori* information about the sensor measurements' distribution. When *a priori* information is lacking, the hypothesis  $H_1$  (signal present in the region of interest) is composite in general, making the design difficult because a UMP detector may not exist [73] [65] and, even if suboptimal designs are adopted, numerical procedures may be required to ensure that a prescribed detection performance is achieved under any distribution within  $H_1$ .

Lack of *a priori* information is inherently present in the design of sensor detection systems to detect a signal generated by an emitter in a random location  $L_e$ . To illustrate this problem, recall the sensor measurement additive model of (2.1), in which the random variable  $A_i$  depends on the distance between the sensor position  $L_i$  and the emitter location  $L_e$ . If the distribution of  $L_e$  and all other random components were known, then the system designer could, at least in theory, use the Bayesian approach [65, p. 198] in order to reduce the composite hypothesis  $H_1$  to a simple hypothesis. However, in many of the applications mentioned in Chapter 1.1, the distribution of the emitter location  $L_e$  is unknown.

One of the approaches to deal with the lack of *a priori* information is to assume a *least favorable distribution* for the random variable whose distribution is unknown [73, p. 84]. The adoption of a least favorable distribution is a conservative and reasonable approach because, if a detector is designed using the least favorable distribution for the random variable whose distribution is unknown, then the detection performance cannot be worse under any other distribution for such a random variable.

In this section, the theory of least favorable distributions is applied to the emitter location distribution. Although the use of least favorable distributions is not new, to the best of the author's knowledge, this is the first time that the theory of least favorable distributions is applied to the emitter location distribution.

Since designs based on a least favorable distribution can be considered too conservative, this section also defines and proposes the use of a most favorable distribution for the emitter location to determine an upper bound for how conservative the design based on a least favorable emitter location distribution is.

## 7.1 The Least Favorable Distribution for the Emitter Location

Assuming that the distribution of the sensor locations  $\{L_i\}_{i=1}^K$ , the measurement noise distribution  $F_W$ , and the amplitude function  $\xi$  are known or given, the only component of the signal model that is unknown is the distribution  $P_{L_e}$  of the emitter location.

For each possible  $P_{L_e}$ , there is a corresponding distribution  $P_{Z_1, \dots, Z_K}$  for the sensor measurements, which means that there is a correspondence between the distributions within the composite hypothesis  $H_1$  and the space of possible distributions  $P_{L_e}$ . Thus, it is possible to address the issue of composite hypothesis by adopting a least favorable distribution for the emitter location  $L_e$ .

In order to precisely define the concept of least favorable distributions for the emitter location, let  $\underline{\phi} := \{\phi_i, P_{Y_i}\}_{i=0}^K$  denote the detection system that uses the fusion function  $\phi_0$ , the sensor functions  $\{\phi_i\}_{i=1}^K$ , and possible randomization variables  $\{Y_i\}_{i=0}^K$ ; and let  $\beta(P_{L_e}, \underline{\phi})$  and  $\alpha(\underline{\phi})$  respectively denote the probability of detection and probability of false alarm obtained by  $\underline{\phi}$  when  $L_e$  is distributed according to  $P_{L_e}$ . Note that the probability of detection depends on other quantities, such as the noise distribution and the distribution of the sensor locations  $\{L_i\}_{i=1}^K$ ;<sup>1</sup> and observe that the probability of false alarm does not depend on the distribution of  $L_e$ .

For any distribution  $P_{L_e}$  of the emitter location, let

$$\underline{\phi}_{P_{L_e}} := \underset{\{\underline{\phi} : \alpha(\underline{\phi}) \leq \alpha_{\max}\}}{\operatorname{argmax}} \beta(P_{L_e}, \underline{\phi}), \quad (7.1)$$

where  $\alpha_{\max} > 0$  is a given desired maximum probability of false alarm;<sup>2</sup> i.e.,  $\underline{\phi}_{P_{L_e}}$  is the Neyman–Pearson detection system when  $P_{L_e}$  is the distribution of  $L_e$ .

Using the definition of Lehmann [73], a distribution  $P_{L_e}^-$  is called *least favorable* if

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}_{P'_{L_e}}) \geq \beta(P_{L_e}^-, \underline{\phi}_{P_{L_e}^-}). \quad (7.2)$$

The definition of Lehmann [73], however, is not enough for design purposes for two reasons:

- Lehmann's definition does not allow one to impose constraints on the detection system. The definition is based on the Neyman–Pearson system, which is a centralized detection system. It is necessary to extend the definition to allow the consideration of distributed detection systems;
- Lehmann's definition does not ensure that  $\underline{\phi}_{P_{L_e}^-}$  will satisfy a minimum performance requirement for all possible distributions. More precisely, if a designer adopts  $\underline{\phi}_{P_{L_e}^-}$  as the detection system, (7.2) may be satisfied while  $\beta(P'_{L_e}, \underline{\phi}_{P_{L_e}^-}) < \beta(P_{L_e}^-, \underline{\phi}_{P_{L_e}^-})$  for some  $P'_{L_e}$ . See Appendix D for an example.

<sup>1</sup>Note that  $\beta(P_{L_e}, \underline{\phi})$  is *not* the probability of detection for a given realization of the sensor locations. More precisely, for a system  $\underline{\phi}$  and a distribution  $P_{L_e}$ , let  $\beta(P_{L_e}, \underline{\phi} | L_1 = l_1, \dots, L_K = l_K)$  denote the probability of detection given the realization  $\{L_1 = l_1, \dots, L_K = l_K\}$ ,  $\beta(P_{L_e}, \underline{\phi}) = \int \beta(P_{L_e}, \underline{\phi} | L_1 = l_1, \dots, L_K = l_K) dP_{L_1, \dots, L_K}(l_1, \dots, l_K)$ .

<sup>2</sup>Although  $\underline{\phi}_{P_{L_e}}$  depends on  $\alpha_{\max}$ ,  $\alpha_{\max}$  is omitted from the notation since a common  $\alpha_{\max}$  is considered whenever comparing two or more detection systems.

These two limitations are addressed below:

### 7.1.1 The Least Favorable Distribution for a Class of Detection Systems

In order to address the first limitation of Lehmann's definition for a least favorable distribution, it is proposed here that the definition of Lehmann be extended as follows: let  $\mathcal{C}$  be a set of detection systems (i.e., a set of possible sensor and fusion functions) satisfying the following condition.

**Condition 7.1:** The set

$$\{\beta(P_{L_e}, \underline{\phi}) : \underline{\phi} \in \mathcal{C}, \alpha(\underline{\phi}) \leq \alpha_{\max}\} \quad (7.3)$$

contains its least upper bound.

For any class  $\mathcal{C}$  satisfying Condition 7.1, define

$$\underline{\phi}_{\{P_{L_e}|\mathcal{C}\}} := \underset{\{\underline{\phi} \in \mathcal{C} : \alpha(\underline{\phi}) \leq \alpha_{\max}\}}{\operatorname{argmax}} \beta(P_{L_e}, \underline{\phi}). \quad (7.4)$$

Although  $\underline{\phi}_{\{P_{L_e}|\mathcal{C}\}}$  is defined just for classes of detection systems  $\mathcal{C}$  satisfying Condition 7.1, this definition is enough for the classes  $\mathcal{C}$  that are considered in this dissertation:

- Let  $\mathcal{C}_c$  denote the class containing all possible sensor and fusion functions. Since  $\mathcal{C}_c$  contains all possible centralized detection systems, it satisfies Condition 7.1 because, for any  $P_{L_e}$ , the centralized detection system with the Neymann–Pearson fusion function provides the maximum probability of detection for a given probability of false alarm.
- Let  $\mathcal{C}_d$  denote the class containing all possible fusion functions and all sensor functions that map into a finite space; i.e.,  $\mathcal{C}_d$  forms the class that contains all possible distributed detection systems. As described in Section 3.1.2.2, there exists an optimum distributed detection system [114, 117]; thus,  $\mathcal{C}_d$  also satisfies Condition 7.1.
- Let  $\mathcal{C}_f$  denote any class containing a finite number of fusion and sensor functions. The class  $\mathcal{C}_f$  satisfies Condition 7.1 because the set defined in (7.3) is finite and therefore contains its least upper bound.
- Let  $\mathcal{C}_s$  denote any class containing a finite number of sensor functions. The class  $\mathcal{C}_s$  satisfies Condition 7.1 because, each possible sensor function and  $P_{L_e}$  induces a unique distribution for the decision random variables at the fusion center, and, for each possible sensor function, the Neyman–Pearson fusion function reaches the maximum possible probability of detection. Since  $\mathcal{C}_s$  contains a finite number of sensor functions, there is a finite number of possible maximum probabilities of detection, and the set defined in (7.3) contains its upper bound.

Given definition (7.4), a distribution  $P_{L_e}^-$  is defined to be a *least favorable distribution for the class of systems  $\mathcal{C}$*  if

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}_{\{P'_{L_e} | \mathcal{C}\}}) \geq \beta(P_{L_e}^-, \underline{\phi}_{\{P_{L_e}^- | \mathcal{C}\}}). \quad (7.5)$$

The extension of the least favorable distribution definition to a class of systems allows the system designer to impose constraints on the set of possible detection systems; however, it does not necessarily address the second limitation of Lehmann's definition; i.e., a least favorable distribution for  $\underline{\phi}_{\{P_{L_e}^- | \mathcal{C}\}}$  does not ensure that  $\underline{\phi}_{\{P_{L_e}^- | \mathcal{C}\}}$  will satisfy a minimum performance requirement for all possible distributions.

In order to address the second limitation of Lehmann's definition, a further specialization of the definition is needed, which will be done in the next subsection.

### 7.1.2 The Least Favorable Distribution for a Detection System

Given that the system designer is interested in adopting a design that ensures the minimum performance requirement for all possible distributions for  $L_e$ , a further specialized definition is proposed here: let  $\underline{\phi}$  be a system that satisfies  $\alpha(\underline{\phi}) \leq \alpha_{\max}$  for some  $\alpha_{\max} > 0$ ; a distribution  $P_{L_e}^-$  is defined to be a *least favorable distribution for the detection system  $\underline{\phi}$*  if

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}) \geq \beta(P_{L_e}^-, \underline{\phi}), \quad (7.6)$$

where it is important to note that the definition is tied to the detection system  $\underline{\phi}$  considered. It is further noted that there might be multiple least favorable distributions for a given detection system  $\underline{\phi}$ .

There are two reasons for adopting this more specialized definition:

- Contrary to the definition of (7.2) or the extension (7.5) to a class of systems, if the detection system  $\underline{\phi}$  satisfies (7.6) and achieves the required detection performance when designed under a least favorable distribution for  $\underline{\phi}$ , then it is ensured that the required detection performance will be achieved under any other distribution of  $L_e$ . A system designer is more interested in ensuring (7.6) than ensuring (7.2) or (7.5).
- The definition of (7.6) is necessary when the system designer is constrained to use a given detection system.

Although the definition of the least favorable distribution for  $\underline{\phi}$  is a special case of the least favorable distribution for class of systems  $\mathcal{C}$  if one considers  $\mathcal{C} = \{\underline{\phi}\}$ , a separate definition is useful because it highlights the performance comparison under a single system.

### 7.1.3 Maximin Detection Systems

Since it is proposed that the system designer ensure the detection performance through (7.6), it is of interest to maximize the right-hand side of (7.6), which motivates the following definition.

A detection system  $\underline{\phi}$  is defined to be *maximin for the class  $\mathcal{C}$*  if it satisfies two conditions:

- for a given  $\alpha_{\max} > 0$ ,  $\alpha(\underline{\phi}) \leq \alpha_{\max}$ ;
- let the distribution  $P_{L_e}$  be a least favorable distribution for  $\underline{\phi}$ ; for any other detection system  $\underline{\phi}^* \in \mathcal{C}$  that satisfies  $\alpha(\underline{\phi}^*) \leq \alpha_{\max}$ ; if  $P_{L_e}''$  is a least favorable distribution for  $\underline{\phi}^*$ , then it satisfies

$$\beta(P_{L_e}, \underline{\phi}) \geq \beta(P_{L_e}'', \underline{\phi}^*); \quad (7.7)$$

which means that  $\underline{\phi}$  and the least favorable distribution  $P_{L_e}$  for  $\underline{\phi}$  maximize the right-hand side of (7.6).

If  $\mathcal{C}$  is the class of all detection systems; i.e.,  $\mathcal{C} = \mathcal{C}_c$ , then  $\underline{\phi}$  is said to be *maximin*; or, when it is important to highlight the type of the system,  $\underline{\phi}$  is said to be the *maximin centralized detection system*. Likewise, if  $\mathcal{C}$  is the class of all distributed detection systems; i.e.,  $\mathcal{C} = \mathcal{C}_d$ , then  $\underline{\phi}$  is said to be *maximin for  $\mathcal{C}_d$* ; or that  $\underline{\phi}$  is the *maximin distributed detection system*.

The next lemma provides sufficient conditions for a system to be maximin for  $\mathcal{C}$ .<sup>3</sup>

**Lemma 7.1.** *Let  $\mathcal{C}$  satisfy Condition 7.1 for a given maximum probability of false alarm  $\alpha_{\max}$ . For a given  $P_{L_e}^-$ , let  $\underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}$  be defined by (7.4). If  $P_{L_e}^-$  satisfies*

$$\forall P_{L_e}', \beta(P_{L_e}', \underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}) \geq \beta(P_{L_e}^-, \underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}); \quad (7.8)$$

*i.e.,  $P_{L_e}^-$  is a least favorable distribution for  $\underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}$ , then the detection system  $\underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}$  is maximin for  $\mathcal{C}$ .*

*Proof:* The first condition for being maximin for  $\mathcal{C}$  is satisfied because a system defined by (7.4) satisfies the probability of false alarm requirement by definition.

To see that the second condition for being maximin for  $\mathcal{C}$  is satisfied, consider any other  $\underline{\phi}^* \in \mathcal{C}$  that satisfies  $\alpha(\underline{\phi}^*) \leq \alpha_{\max}$  and a distribution  $P_{L_e}''$  such that

$$\forall P_{L_e}', \beta(P_{L_e}', \underline{\phi}^*) \geq \beta(P_{L_e}'', \underline{\phi}^*); \quad (7.9)$$

i.e.,  $P_{L_e}''$  is a least favorable distribution for  $\underline{\phi}^*$ , and write

$$\beta(P_{L_e}^-, \underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}) \geq \beta(P_{L_e}^-, \underline{\phi}^*) \geq \beta(P_{L_e}'', \underline{\phi}^*), \quad (7.10)$$

where the first inequality follows from the optimality of  $\underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}$  when  $P_{L_e}^-$  is the distribution of  $L_e$  and the second inequality follows from (7.9).<sup>4</sup> ■

<sup>3</sup>The definition of maximin and Lemma 7.1 are similar to and were significantly influenced by [73, p. 320]. The modifications adopted in this dissertation facilitate the understanding and are sufficient for the setting being considered.

<sup>4</sup>It is interesting to observe that, in this case,  $P_{L_e}^-$  is not only a least favorable distribution for the detection system  $\underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}$ , but also a least favorable distribution that satisfies (7.5) because [73]

$$\forall P_{L_e}', \beta(P_{L_e}', \underline{\phi}_{\{P_{L_e}'|\mathcal{C}\}}) \geq \beta(P_{L_e}', \underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}) \geq \beta(P_{L_e}^-, \underline{\phi}_{\{P_{L_e}^-|\mathcal{C}\}}), \quad (7.11)$$

where the first inequality is justified because  $\underline{\phi}_{\{P_{L_e}'|\mathcal{C}\}}$  is the optimal detection system when considering  $L_e$  distributed with  $P_{L_e}'$ .

It is important to mention that, instead of considering that  $H_1$  corresponds to  $\theta = \theta_1$  for a given  $\theta_1 > 0$  and is therefore a simple hypothesis, Lehmann defines the concept of a least favorable distribution and the maximin detection system when  $H_1$  is composite due to randomness in  $\theta$  [73]. The results presented in this dissertation illustrate that the definition of a least favorable distribution and maximin detection systems can be generalized to any random variable that causes  $H_1$  to be composite.

Before concluding this section, note that a least favorable distribution  $P_{L_e}$  for a detection system  $\underline{\phi}$  depends not only on  $\underline{\phi}$ , but also on any other distribution that influences the measurements. For instance, consider a system  $\underline{\phi}$  in which sensors' locations are distributed according to a first distribution  $P_{L_i}^{(1)}$ , and assume that  $P_{L_e}^{(1)}$  is a least favorable distribution for  $\underline{\phi}$ . If the sensors' locations are instead distributed according to a second distribution  $P_{L_i}^{(2)}$ , a different distribution  $P_{L_e}^{(2)}$  may be a least favorable distribution for  $\underline{\phi}$ .

## 7.2 The Most Favorable Distribution for a Detection System

In spite of simplifying the design and ensuring a detection performance, the use of least favorable distributions may be criticized because it may lead to a design that is too conservative.

In order to investigate how conservative the approach of least favorable distributions is, its performance will be compared against the performance achieved by the *most favorable distribution*.

Using the notation introduced in Section 7.1, when  $\mathcal{C}$  satisfies Condition 7.1, a distribution  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$  is a *most favorable distribution for the class of systems  $\mathcal{C}$*  if

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}_{\{P'_{L_e}|\mathcal{C}\}}) \leq \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}); \quad (7.12)$$

and, for any system  $\underline{\phi}$  that satisfies  $\alpha(\underline{\phi}) \leq \alpha_{\max}$  for some  $\alpha_{\max} > 0$ , a distribution  $P_{L_e}^+$  is a *most favorable distribution for the detection system  $\underline{\phi}$*  if

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}) \leq \beta(P_{L_e}^+, \underline{\phi}), \quad (7.13)$$

where, as in the definition of a least favorable distribution for a detection system, it is noted that the definition is tied to the detection system  $\underline{\phi}$  considered, and there might be multiple most favorable distributions for a given detection system.

The definition of a most favorable distribution for a detection system is of interest when a system designer considers a particular detection system  $\underline{\phi}$  and adopts a least favorable distribution for  $\underline{\phi}$  because a most favorable distribution for  $\underline{\phi}$  provides an upper bound for the detection performance of a system under any distribution; thus, when using both a least and a most favorable distributions, a system designer can determine a range for the detection performance of a system as the emitter location distribution changes.

More precisely, if the system designer chooses a detection system  $\underline{\phi}$  that satisfies  $\alpha(\underline{\phi}) \leq \alpha_{\max}$  and is able to determine a least favorable distribution  $P_{L_e}^-$  and a most favorable distribution  $P_{L_e}^+$  for such a system, then the performance

of the detection system  $\underline{\phi}$  under any distribution  $P_{L_e}$  is guaranteed to satisfy

$$\beta(P_{L_e}^+, \underline{\phi}) \geq \beta(P_{L_e}, \underline{\phi}) \geq \beta(P_{L_e}^-, \underline{\phi}). \quad (7.14)$$

### 7.2.1 Maximax Detection Systems

It is also of interest to find the highest upper bound for the detection performance as one varies the detection system. In order to find such upper bound, a detection system  $\underline{\phi}$  is defined to be *maximax for the class  $\mathcal{C}$*  if it satisfies two conditions:

- for a given  $\alpha_{\max} > 0$ ,  $\alpha(\underline{\phi}) \leq \alpha_{\max}$ ;
- let the distribution  $P_{L_e}$  be a most favorable distribution for  $\underline{\phi}$ ; for any other detection system  $\underline{\phi}^* \in \mathcal{C}$  that satisfies  $\alpha(\underline{\phi}^*) \leq \alpha_{\max}$  and any other distribution  $P'_{L_e}$ ,

$$\beta(P_{L_e}, \underline{\phi}) \geq \beta(P'_{L_e}, \underline{\phi}^*); \quad (7.15)$$

which means that  $\underline{\phi}$  and the most favorable distribution  $P_{L_e}$  for  $\underline{\phi}$  maximize the right-hand side of (7.13). If  $\mathcal{C}$  is the class of all detection systems,  $\underline{\phi}$  is said to be *maximax*.

The next lemma provides sufficient conditions for a system to be maximax for a class  $\mathcal{C}$ .

**Lemma 7.2.** *Let  $\mathcal{C}$  satisfy Condition 7.1 for a given maximum probability of false alarm  $\alpha_{\max}$ . For a given  $P_{L_e}^+$ , let  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$  be defined by (7.4). If  $P_{L_e}^+$  satisfies*

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}_{\{P'_{L_e}|\mathcal{C}\}}) \leq \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}); \quad (7.16)$$

*then the detection system  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$  is maximax for  $\mathcal{C}$ .*

*Proof:* By definition, the system  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$  satisfies the first condition for being maximax for  $\mathcal{C}$ .

Before showing that the second condition for being maximax for  $\mathcal{C}$  is satisfied, it will first be shown that  $P_{L_e}^+$  is a most favorable distribution for  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$ . To see this, observe that

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}_{\{P'_{L_e}|\mathcal{C}\}}) \leq \beta(P'_{L_e}, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}) \leq \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}), \quad (7.17)$$

where (7.4) is used to justify the first inequality, and (7.16) is used to justify the second inequality. This means that  $P_{L_e}^+$  satisfies (7.13) for  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$ ; i.e.,  $P_{L_e}^+$  is a most favorable distribution for the system  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$ .

It is now shown that (7.15) is satisfied. For this, pick any  $\underline{\phi}^* \in \mathcal{C}$  that satisfies  $\alpha(\underline{\phi}^*) \leq \alpha_{\max}$  and any distribution  $P'_{L_e}$  and use the same rationale used to reach (7.17) in order to write

$$\forall P'_{L_e}, \beta(P'_{L_e}, \underline{\phi}^*) \leq \beta(P'_{L_e}, \underline{\phi}_{\{P'_{L_e}|\mathcal{C}\}}) \leq \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}). \quad (7.18)$$

■



### 7.3 Sufficient Conditions for Least and Most Favorable Distributions when Sensors Locations are Random Variables

In this section, it will be shown that there are sufficient conditions under which the distribution  $P_{L_e}^-$  that satisfies  $P[L_e \in S_e^-] = 1$ , where  $S_e^-$  satisfies the conditions of the set  $S_e^*$  in either one of the Propositions 6.2 or 6.3, and the distribution  $P_{L_e}^+$  that satisfies  $P[L_e \in S_e^+] = 1$  for  $S_e^+ = \{0\}$  are respectively a least and a most favorable distributions for certain detection systems of interest.

More precisely, sufficient conditions are presented for detection systems that belong to either the class  $\mathcal{D}_c$  of centralized detection systems or the class  $\mathcal{D}_d$  of distributed detection systems.

**Definition of  $\mathcal{D}_c$ :** A centralized detection system  $\underline{\phi} := (\phi_0, P_{Y_0})$  belongs to  $\mathcal{D}_c$  if the communication subsystem offers  $K$  parallel and error-free communication channels, and the fusion function  $\phi_0(\mathbf{z}, y)$  is given by

$$\phi_0(\mathbf{z}, y) = \mathbf{1} \left\{ \prod_{i=1}^K T_{0,i}(z_i) \in I_{0,y} \right\}, \quad (7.19)$$

where  $\mathbf{z} := (z_1, \dots, z_K)$ ,  $y$  is the realization of the randomization random variable  $Y_0 \in \{1, 2\}$ ,  $I_{0,1} := (t_0, \infty]$  and  $I_{0,2} := [t_0, \infty]$  for a given  $t_0$ , and  $T_{0,i}(z)$  is a nonnegative and nondecreasing function on  $\mathcal{Z}$ ; i.e.,

$$\forall i \in \{1, \dots, K\}, \forall z_1, z_2 \in \mathcal{Z}, z_1 \leq z_2 \Rightarrow T_{0,i}(z_1) \leq T_{0,i}(z_2), \quad (7.20)$$

where  $\mathcal{Z}$  is the space in which any  $Z_i$  assumes its values.

The class  $\mathcal{D}_c$  of detection systems is general enough to include many centralized detection systems of interest, including practical and optimal systems, as highlighted by the following propositions:

**Proposition 7.3.** *The deterministic centralized detection system with  $K$  parallel and error-free communication channels that uses the fusion function*

$$\phi_0^{(\Sigma)}(\mathbf{z}, y_0) = \mathbf{1} \left\{ \sum_{i=1}^K z_i > t_0 \right\} \quad (7.21)$$

*is an element of  $\mathcal{D}_c$ .*

*Proof:* Just observe that (7.21) is of the form specified by (7.19) if  $T_{0,i}(z_i) = \exp\{z_i\}$  and such a  $T_{0,i}(z_i)$  is a nonnegative and nondecreasing real function. ■

**Proposition 7.4.** *Assume that the measurement  $Z_i$  follows the additive model of (2.1); i.e.,  $Z_i = \mathbf{1}\{\theta > 0\} \cdot A_i + W_i$  and adopt Assumptions I, II, III, V, and VIII of Chapter 5. Let  $Z_i$  assume values in some  $\mathcal{Z} \subset \mathbb{R}$  and let  $\mathcal{Z}_+ := \mathcal{Z} \cap \mathbb{R}_+$ . Let  $A_i$  assume values in  $\mathcal{Z}_+$  and  $W_i$  assume values in  $\mathcal{Z}$ . Assume  $f_W(w) > 0$  for any  $w \in \mathcal{Z}$  and consider that*

$$\forall w_1 \leq w_2 \in \mathcal{Z}, \forall a \in \mathcal{Z}_+, \frac{f_W(w_1 - a)}{f_W(w_1)} \leq \frac{f_W(w_2 - a)}{f_W(w_2)}, \quad (7.22)$$

i.e.,  $f_W(w-a)/f_W(w)$  is nondecreasing in  $w \in \mathcal{Z}$  for all  $a \in \mathcal{Z}_+$ . Pick any subset  $S'_e$  of the region of interest  $S_e$  for which  $P[L_e \in S'_e] > 0$ .

The centralized detection system with  $K$  parallel and error-free communication channels and fusion function given by

$$\phi_0(\mathbf{z}, y) := \mathbf{1} \left\{ \prod_{i=1}^K \mathcal{L}_{Z_i, S'_e}(z_i) \in I_{0,y} \right\}, \quad (7.23)$$

where  $\{I_{0,y}\}_{y=1}^2$  are intervals defined by  $(t_{0,y}, \infty]$  or  $[t_{0,y}, \infty]$  for a certain  $\{t_{0,y}\}_{y=1}^2$ , and

$$\mathcal{L}_{Z_i, S'_e}(z_i) := \begin{cases} \frac{f_{Z_i|\theta=\theta_1, L_e \in S'_e}(z_i)}{f_{Z_i|\theta=0}(z_i)}, & f_{Z_i|\theta=0}(z_i) > 0, \\ +\infty, & f_{Z_i|\theta=0}(z_i) = 0. \end{cases} \quad (7.24)$$

is an element of  $\mathcal{D}_c$ .

*Proof:* To see that the fusion function (7.23) is on the form specified by (7.19), let  $T_{0,i}(z_i, y) = \mathcal{L}_{Z_i, S'_e}(z_i)$ . Such  $T_{0,i}$  is nonnegative and to see that it also satisfies (7.20); i.e., it is a nondecreasing function of  $z_i$  on  $\mathcal{Z}$ , observe that, from Lemma E.1 in Appendix E,  $f_{Z_i|\theta=\theta_1, L_e \in S'_e}(z_i) = \int_{\mathcal{Z}_+} f_W(z_i - a) dP_{A_i|\theta=\theta_1, L_e \in S'_e}(a)$  and  $f_{Z_i|\theta=0}(z_i) = f_W(z_i)$ , which means that  $\mathcal{L}_{Z_i, S'_e}(z) = \int_{\mathcal{Z}_+} f_W(z - a)/f_W(z) dP_{A_i|\theta=\theta_1, L_e \in S'_e}(a)$  for any  $z \in \mathcal{Z}$  since it is assumed that  $f_W(w) > 0$  for all  $w \in \mathcal{Z}$ . Since Assumption V gives that  $A_i \geq 0$  and since (7.22) is satisfied for all  $a \in \mathcal{Z}_+$ , for any  $z_1 \leq z_2 \in \mathcal{Z}$ ,

$$\mathcal{L}_{Z_i, S'_e}(z_1) = \int_{\mathcal{Z}_+} \frac{f_W(z_1 - a)}{f_W(z_1)} dP_{A_i|\theta=\theta_1, L_e \in S'_e}(a) \leq \int_{\mathcal{Z}_+} \frac{f_W(z_2 - a)}{f_W(z_2)} dP_{A_i|\theta=\theta_1, L_e \in S'_e}(a) = \mathcal{L}_{Z_i, S'_e}(z_2). \quad (7.25)$$

■

While  $\mathcal{D}_c$  is general enough for centralized detection systems, a more elaborate class is required for distributed detection systems; not only because they use non-trivial sensor functions, but also because multiple sensor functions can be considered. As described in Section 3.1.2.2, distributed detection systems may involve the dependent randomization among different sets of fusion and sensor functions. More precisely, a distributed detection system may operate as follows: before a decision, a realization  $y$  is drawn from a discrete random variable  $Y$  that assumes values in  $\{1, \dots, M\}$ ; and the value of  $y$  dictates that each sensor is to use a particular sensor function  $\phi_i(z_i, y)$  to treat its particular measurement  $z_i$ , and the fusion function is to use a particular fusion function  $\phi_0(\mathbf{u}, y)$  to treat the vector of sensor outputs  $\mathbf{u}$ . The need to model such types of distributed detection systems motivate the following definition:

**Definition of  $\mathcal{D}_d$ :** A distributed detection system  $\underline{\phi} := (\{\phi_i, P_{Y_i}\}_{i=1}^K)$  belongs to  $\mathcal{D}_d$  if the communication subsystem offers  $K$  parallel and error-free communication channels, and

- (i) the various random variables  $Y_i$  are equal almost surely to a common random variable  $Y \in \{1, \dots, M\}$ ; and  $Y$  is independent of  $\mathbf{Z}$  and independent of any other random variable to which  $\mathbf{Z}$  is dependent;
- (ii) the sensor functions are given by

$$\phi_i(z_i, y) = g_i(T_i(z_i), y), \quad (7.26)$$

where  $g_i(x, y)$  is a nondecreasing real function of  $x$  for any  $y$ , and  $T_i : \mathcal{Z} \rightarrow \mathbb{R}$  is a nondecreasing function on  $\mathcal{Z}$ ; i.e.,

$$\forall i, \forall z_1, z_2 \in \mathcal{Z}, z_1 \leq z_2 \Rightarrow T_i(z_1) \leq T_i(z_2), \quad (7.27)$$

where  $\mathcal{Z}$  is the space in which any  $Z_i$  assumes its values;

(iii) the fusion function  $\phi_0(\mathbf{u}, y)$  is given by

$$\phi_0(\mathbf{u}, y) = \mathbf{1} \left\{ \prod_{i=1}^K T_{0,i}(u_i, y) \in I_{0,y} \right\}, \quad (7.28)$$

where  $y$  is the realization of the randomization random variable  $Y$ ,  $\mathbf{u} := (u_1, \dots, u_K)$  with  $u_i = \phi_i(z_i, y)$ ,  $\{I_{0,y}\}_{y=1}^M$  are intervals defined by  $(t_{0,y}, \infty]$  for given  $\{t_{0,y}\}_{y=1}^M$ ; <sup>5</sup> and for any  $i \in \{1, \dots, K\}$  and any  $y \in \{1, \dots, M\}$ ,  $T_{0,i}(u, y)$  is a nonnegative and nondecreasing function on  $u \in \mathcal{U}_{i,y} := \{0, \dots, U_{i,y}^{\max}\}$  for given  $\{\{U_{i,y}^{\max}\}_{y=1}^M\}_{i=1}^K$ ; i.e.,

$$\forall i \in \{1, \dots, K\}, \forall y \in \{1, \dots, M\}, \forall u_1, u_2 \in \mathcal{U}_{i,y}, u_1 \leq u_2 \Rightarrow 0 \leq T_{0,i}(u_1, y) \leq T_{0,i}(u_2, y). \quad (7.29)$$

The following propositions illustrate that several of the distributed detection systems of interest fit into the definition of the class  $\mathcal{D}_d$ .

**Proposition 7.5.** *The deterministic distributed detection system with  $K$  parallel and error-free communication channels that uses the following fusion and sensor functions*

$$\phi_0^{(\Sigma)}(\mathbf{u}, y_0) = \mathbf{1} \left\{ \sum_{i=1}^K u_i > t_0 \right\}, \quad (7.30)$$

$$u_i := \phi_i^{(\Sigma)}(z_i, y_i) = \sum_{u=0}^{U^{\max}} u \cdot \mathbf{1}\{t_{i,u} < z_i \leq t_{i,u+1}\}, \quad (7.31)$$

for some  $-\infty = t_{i,0} < \dots < t_{i,U^{\max}} < t_{i,U^{\max}+1} = \infty$ , is an element of  $\mathcal{D}_d$ . Note that, when  $U^{\max} = 1$ , the usual fusion functions ‘OR’, ‘AND’, and ‘MAJORITY’ can be obtained at different values of  $t_0$ .

*Proof:* Condition (i) for being in  $\mathcal{D}_d$  is satisfied because being deterministic is equivalent to assuming randomization random variables  $\{Y_i\}_{i=1}^K$  that are degenerative and independent of  $\{Z_i\}_{i=1}^K$ ,  $\{L_i\}_{i=1}^K$ ,  $L_e$ ,  $\{W_i\}_{i=1}^K$ , and  $\theta$ . To see why condition (ii) for being in  $\mathcal{D}_d$  is satisfied, note that  $\phi_i^{(\Sigma)}(z_i, y_i)$  has the form of (7.26) if  $g_i(x, y) = \sum_{u=0}^{U^{\max}} u \cdot \mathbf{1}\{t_{i,u} < x \leq t_{i,u+1}\}$  and  $T_i(z) = z$  and  $g_i(x, y)$  is a nondecreasing step function because  $-\infty = t_{i,0} < \dots < t_{i,U^{\max}} < t_{i,U^{\max}+1} = \infty$ . Condition (iii) for being in  $\mathcal{D}_d$  is satisfied for the same reason as in Proposition 7.3. ■

**Proposition 7.6.** *Assume that the measurement  $Z_i$  follows the additive model of (2.1); i.e.,  $Z_i = \mathbf{1}\{\theta > 0\} \cdot A_i + W_i$  and adopt Assumptions I, II, III, V, and VIII of Chapter 5.*

<sup>5</sup>Considering  $I_{0,y} = (t_{0,y}, \infty]$  is enough for the fusion function of a distributed detection system because there is a finite number of possible realizations of  $\{U_i\}_{i=1}^K$ , which imply a finite number of possible values for the product  $\prod_{i=1}^K T_{0,i}(u_i, y)$ . Therefore, it is possible to use intervals of the form  $(t_{0,y}, \infty]$  to build the decision regions.

Let  $Z_i$  assume values in some  $\mathcal{Z} \subset \mathbb{R}$  and let  $\mathcal{Z}_+ := \mathcal{Z} \cap \mathbb{R}_+$ . Let  $A_i$  assume values in  $\mathcal{Z}_+$  and  $W_i$  assume values in  $\mathcal{Z}$ . Assume  $f_W(w) > 0$  for any  $w \in \mathcal{Z}$  and assume that

$$\forall w_1 \leq w_2 \in \mathcal{Z}, \forall a \in \mathcal{Z}_+, \frac{f_W(w_1 - a)}{f_W(w_1)} \leq \frac{f_W(w_2 - a)}{f_W(w_2)}, \quad (7.32)$$

i.e.,  $f_W(w - a)/f_W(w)$  is nondecreasing in  $w \in \mathcal{Z}$  for all  $a \in \mathcal{Z}_+$ . Pick any subset  $S'_e$  of the region of interest  $S_e$  for which  $P[L_e \in S'_e] > 0$ . Consider the distributed detection system  $\underline{\phi}$  with  $K$  parallel and error-free communication channels, with dependent randomization; i.e.,  $Y_0 = Y_1 = \dots = Y_K = Y \in \{1, \dots, M\}$  almost surely, where  $P[Y = m] > 0$  for all  $m \in \{1, \dots, M\}$ .

Assume that  $\underline{\phi}$  has sensor functions given by

$$u_i := \phi_i(z_i, y) = \sum_{u=0}^{U_{i,y}^{\max}} u \cdot \mathbf{1}\left\{\mathcal{L}_{Z_i, S'_e}(z_i) \in I_{i,y,u}\right\}, \quad (7.33)$$

where  $\mathcal{L}_{Z_i, S'_e}(z_i)$  is defined as in (7.24); and each  $I_{i,y,u}$  is an interval defined by thresholds  $t_{i,y,u}$  and  $t_{i,y,u+1}$ . Assume that, for any  $i$  and  $y$ ,

- The thresholds  $\{t_{i,y,u}\}_{u=0}^{U_{i,y}^{\max}}$  satisfy

$$0 = t_{i,y,0} \leq t_{i,y,1} \leq \dots \leq t_{i,y,U_{i,y}^{\max}} \leq t_{i,y,U_{i,y}^{\max}+1} = \infty; \quad (7.34)$$

- The intervals  $\{I_{i,y,u}\}_{u=0}^{U_{i,y}^{\max}}$  are disjoint and their union equals  $\mathbb{R}_+ \cup \{\infty\}$ ; and
- The intervals  $\{I_{i,y,u}\}_{u=0}^{U_{i,y}^{\max}}$  are such that<sup>6</sup>

$$\forall u \in \{0, 1, \dots, U_{i,y}^{\max}\}, P[U_i = u | \theta = 0, Y = y] > 0. \quad (7.35)$$

Suppose that  $\underline{\phi}$  has a fusion function given by

$$\phi_0(\mathbf{u}, y) := \mathbf{1}\left\{\prod_{i=1}^K \mathcal{L}_{U_i, S'_{e,y}}(u_i) \in I_{0,y}\right\}, \quad (7.36)$$

where  $\{I_{0,y}\}_{y=1}^M$  are intervals defined by  $(t_{0,y}, \infty]$  for a certain  $\{t_{0,y}\}_{y=1}^M$ ; and

$$\mathcal{L}_{U_i, S'_{e,y}}(u_i) := \begin{cases} \frac{P_{U_i|\theta=\theta_1, L_e \in S'_{e,y}=y}(u_i)}{P_{U_i|\theta=0, Y=y}(u_i)}, & P_{U_i|\theta=0, Y=y}(u_i) > 0, \\ +\infty, & P_{U_i|\theta=0, Y=y}(u_i) = 0. \end{cases} \quad (7.37)$$

Then the distributed detection system  $\underline{\phi}$  is an element of  $\mathcal{D}_d$ .

<sup>6</sup>This assumption is not restrictive because, as argued in [123] and shown in details in Lemma F.1 in the appendix, if the thresholds were such that  $\exists u^- \in \{0, 1, \dots, U_{i,y}^{\max}\}$  with  $P[U_i = u^- | \theta = 0, Y = y] = 0$ ; then it follows that  $P[U_i = u^- | \theta = \theta_1, Y = y] = 0$  as well and it is possible to change the sensor function  $\phi_i$  by excluding this sensor output; i.e., by reducing  $U_{i,y}^{\max}$  by 1, changing the labels of the sensor outputs, and making the appropriate changes in the fusion function, such that condition (7.35) is satisfied without changing the probability of detection or the probability of false alarm at the fusion center.

*Proof:* Condition (i) for being in  $\mathcal{D}_d$  is satisfied by definition. To show that condition (ii) for being in  $\mathcal{D}_d$  is satisfied, let  $g_i(x, y) = \sum_{u=0}^{U_{i,y}^{\max}} u \cdot \mathbf{1}\{x \in I_{i,y,u}\}$  and  $T_i(z) = \mathcal{L}_{Z_i|S_e'}(z)$ . Because of (7.34),  $g_i(x, y)$  is a nondecreasing function for any  $y$  and  $T_i(z)$  is a nondecreasing function on  $\mathcal{Z}$  as shown in Proposition 7.4.

To show that condition (iii) for being in  $\mathcal{D}_d$  is satisfied, observe that the fusion function (7.36) is on the form specified by condition (iii) with  $T_{0,i}(u_i, y) = \mathcal{L}_{U_i|S_e', Y}(u_i)$ . Such a  $T_{0,i}(u_i, y)$  is nonnegative. To see that  $T_{0,i}(u_i, y)$  satisfies (7.29), observe that for either  $u^-$  or  $u^+$ ,

$$T_{0,i}(u, y) = \frac{P_{U_i|\theta=\theta_1, L_e \in S_e', Y=y}(u)}{P_{U_i|\theta=0, Y=y}(u)} \quad (7.38)$$

and recall that, for either  $j = 0$  or  $j = \theta_1$ ,

$$\begin{aligned} P_{U_i|\theta=j, Y=y, L_e \in S_e'}(u) &= P[\phi_i(Z_i, y) = u | \theta = j, Y = y, L_e \in S_e'] \\ &= P[T_i(Z_i) \in I_{i,y,u} | \theta = j, Y = y, L_e \in S_e']. \end{aligned} \quad (7.39)$$

Since the thresholds that define  $I_{i,y,u^-}$  and  $I_{i,y,u^+}$  satisfy  $t_{i,y,u^-} \leq t_{i,y,u^-+1} \leq t_{i,y,u^+} \leq t_{i,y,u^++1}$ , it is possible to use Lemma 3 of [123] to reach that

$$\begin{aligned} T_{0,i}(u^-, y) &= \frac{P[T_i(Z_i) \in I_{i,y,u^-} | \theta = \theta_1, L_e \in S_e']}{P[T_i(Z_i) \in I_{i,y,u^-} | \theta = 0]} \\ &\leq \frac{P[T_i(Z_i) \in I_{i,y,u^+} | \theta = \theta_1, L_e \in S_e']}{P[T_i(Z_i) \in I_{i,y,u^+} | \theta = 0]} = T_{0,i}(u^+, y). \end{aligned} \quad (7.40)$$

For completeness, the steps taken in Lemma 3 of [123] are presented in the context of this proposition: note that  $P[T_i(Z_i) \in I_{i,y,u^-} | \theta = \theta_1, L_e \in S_e'] = \int_{\{T_i(z) \in I_{i,y,u^-}\}} f_{Z_i|\theta=\theta_1, L_e \in S_e'}(z) d\mu(z)$ . Since  $I_{i,y,u^-}$  is bounded for the  $u^-$  being considered,  $T_i(z)$  is finite, which implies  $f_{Z_i|\theta=\theta_0}(z) > 0$ . Multiplying and dividing the integrand by  $f_{Z_i|\theta=\theta_0}(z)$  and recognizing  $T_i(z)$ , the integrand becomes  $\int_{\{T_i(z) \in I_{i,y,u^-}\}} T_i(z) f_{Z_i|\theta=\theta_0}(z) d\mu(z)$ . Bounding this integral by upper and lower bounds on  $T_i(z)$ , one reaches that the l.h.s. must be in between  $t_{i,y,u^-}$  and  $t_{i,y,u^-+1}$ . Similarly, the r.h.s. of the inequality in (7.40) must be in between  $t_{i,y,u^+}$  and  $t_{i,y,u^++1}$  and the inequality follows because  $t_{i,y,u^-+1} \leq t_{i,y,u^+}$ . ■

It is now shown that least and most favorable distributions for detection systems that belong to either  $\mathcal{D}_c$  or  $\mathcal{D}_d$  can be found when the region of interest  $S_e$ , the emitter location  $L_e$ , and the sensor locations  $\{L_i\}_{i=1}^K$  satisfy the conditions of Propositions 6.2 or 6.3; i.e., when  $S_e$  is a ball and the emitter location  $L_e$  is distributed on the boundary of this ball, or when  $S_e$  is a regular convex polygon and  $L_e$  is distributed on the set of the polygon's vertices, and the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution on the deployment region  $S_s$ .

**Proposition 7.7.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. For each  $i$ , let  $Z_i$  assume values in  $\mathcal{Z}$ , with  $\mathcal{Z} = \mathbb{R}$ ,  $\mathcal{Z} = \mathbb{R}_+$ ,  $\mathcal{Z} = \mathbb{Z}$ , or  $\mathcal{Z} = \mathbb{Z}_+$ . Let  $\mathcal{Z}_+ = \mathcal{Z} \cap \mathbb{R}_+$ . Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , and define  $S_e^- := S_e \cap \partial B_{r_M}(0)$  and  $S_e^+ := \{0\}$ . If*

1. the sensor detection system belongs to either the class  $\mathcal{D}_c$  or  $\mathcal{D}_d$  of detection systems;
2. the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  or in  $S_s = c \cdot B_{r_M}(0)$  for some  $c > 0$ ;
3. for any  $a_i$ ,  $P[A_i > a_i | \theta = \theta_1, \xi(\mathbf{dist}(L_i, L_e)) = x]$  is a nondecreasing function of  $x$ ;

then

- the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  when the distribution of  $L_e$  satisfies either  $P[L_e \in S_e^-] = 1$  or  $P[L_e \in S_e^+] = 1$ ;
- a distribution for  $L_e$  such that  $P[L_e \in S_e^-] = 1$  is a least favorable distribution for the detection system; and
- the distribution for  $L_e$  that satisfies  $P[L_e \in S_e^+] = 1$  is a most favorable distribution for the detection system.

*Proof:* The first conclusion follows from Propositions 6.2 and 6.3. To see this, observe that, when  $S_e = B_{r_M}(0)$ , both  $S_e^- = \partial B_{r_M}(0)$  and  $S_e^+ = \{0\}$  satisfy the condition on  $S_e^*$  of Proposition 6.2 when  $r = r_M$  and  $r = 0$  respectively. When  $S_e$  is a regular convex polygon, both  $S_e^- = S_e \cap \partial B_{r_M}(0)$  and  $S_e^+ = \{0\}$  satisfy the condition on  $S_e^*$  of Proposition 6.3 when  $c_1 = 1$  and  $c_1 = 0$  respectively.

Suppose that the detection system belongs to  $\mathcal{D}_d$ . To establish the second conclusion, it is necessary to show that  $P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1]$  under any  $P_{L_e}$  is bounded below by  $P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1]$  under  $P_{L_e^-}$ . Since Propositions 6.2 and 6.3 show that  $\{Z_i\}_{i=1}^K$  and  $L_e$  are independent under any  $P_{L_e^-}$  satisfying  $P[L_e \in S_e^-] = 1$ , it is enough to consider that  $P_{L_e^-}$  satisfies  $P[L_e = l_e^-] = 1$  for any  $l_e^- \in S_e^-$ . Thus, it suffices to show that

$$\int P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1, L_e = l_e] dP_{L_e}(l_e) \geq P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1, L_e = l_e^-]. \quad (7.41)$$

Observe that the probability of false alarm is constant for any  $P_{L_e}$ .

To prove (7.41), it is enough to show that, for all  $l_e$ ,

$$P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1, L_e = l_e] \geq P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1, L_e = l_e^-]. \quad (7.42)$$

Since any detection system in  $\mathcal{D}_d$  has a common randomization random variable  $Y$  independent of  $\{Z_i\}_{i=1}^K$  and independent of any random variable on which  $\{Z_i\}_{i=1}^K$  depends, which includes  $L_e, \forall l_e$ ,

$$P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_1, L_e = l_e] = \sum_{m=1}^M P[\phi_0(\mathbf{U}, m) = 1 | \theta = \theta_1, L_e = l_e, Y = m] P[Y = m]. \quad (7.43)$$

From the form of the fusion function, (7.42) follows if,  $\forall m$ ,

$$P\left[\prod_{i=1}^K T_{0,i}(U_i, m) \in I_{0,m} | \theta = \theta_1, L_e = l_e, Y = m\right] \geq P\left[\prod_{i=1}^K T_{0,i}(U_i, m) \in I_{0,m} | \theta = \theta_1, L_e = l_e^-, Y = m\right]. \quad (7.44)$$

Since  $\{T_{0,i}\}_{i=1}^K$  are nonnegative for any  $y$  and  $I_{0,m} = (t_{0,m}, \infty]$ , (7.44) follows if

$$\forall t, P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e, Y = m] \geq P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e^-, Y = m], \quad (7.45)$$

as proven in Lemma F.2 in the appendix. Note that, once (7.45) is proven, (7.44) also follows for intervals  $I_{0,m} = [t_{0,m}, \infty]$ , as shown in Lemma F.3 in the appendix.

To show (7.45), observe that, for any  $l_e$  and any  $t$ ,

$$P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e, Y = m] = \sum_{u=0}^{U_{i,m}^{\max}} \mathbf{1}\{T_{0,i}(u, m) > t\} \cdot P[U_i = u | \theta = \theta_1, L_e = l_e, Y = m] \quad (7.46)$$

$$= \sum_{u=u_t}^{U_{i,m}^{\max}} P[U_i = u | \theta = \theta_1, L_e = l_e, Y = m] = P[U_i \geq u_t | \theta = \theta_1, L_e = l_e, Y = m], \quad (7.47)$$

where  $u_t$  is such that  $T_{0,i}(u_t, m) > t$  and, if such a  $u_t > 0$ ,  $T_{0,i}(u_t - 1, m) \leq t$ ; and (7.47) follows because any detection system that belongs to  $\mathcal{D}_d$  satisfies (7.29); i.e.,  $\forall u > u_t$ ,  $T_{0,i}(u, m) \geq T_{0,i}(u_t, m) > t$ .

Since  $U_i = \phi_i(Z_i, m)$ , which is assumed to be a nondecreasing function of  $T_i(z)$ ,

$$\begin{aligned} P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e, Y = m] &= P[\phi_i(Z_i, m) \geq u_t | \theta = \theta_1, L_e = l_e] \\ &= P[T_i(Z_i) \in J_u | \theta = \theta_1, L_e = l_e], \end{aligned} \quad (7.48)$$

where  $J_u = [t_u, \infty]$  or  $J_u = (t_u, \infty]$  for some  $t_u$ ; and the conditioning on  $Y = m$  can be dropped since  $Z_i$  and  $Y$  are independent. Thus, (7.45) will follow if,  $\forall t_u$ ,

$$P[T_i(Z_i) > t_u | \theta = \theta_1, L_e = l_e] \geq P[T_i(Z_i) > t_u | \theta = \theta_1, L_e = l_e^-], \quad (7.49)$$

and the condition for  $J_u = [t_u, \infty]$  follows from Proposition F.3 in the appendix.

To prove (7.49), observe that  $P[T_i(Z_i) > t_u | \theta = \theta_1, A_i = a, L_e = l_e]$  as a function of  $l_e$  remains constant and equal to  $P[T_i(Z_i) > t_u | \theta = \theta_1, A_i = a]$  for any  $l_e \in S_e$  because  $T_i(Z_i)$  depends on  $L_e$  through  $A_i$ . For details, see the *proof* of Lemma E.1 in the appendix. Observe further that, since it is assumed in Assumption V that  $A_i \geq 0$ , and since any realization of  $Z_i$  belongs to  $\mathcal{Z}_+$ , any realization of  $A_i$  must belong to  $\mathcal{Z}_+$ .

It then follows that (7.49) is equivalent to

$$\begin{aligned} &\int_{\mathcal{Z}_+} P[T_i(Z_i) > t_u | \theta = \theta_1, A_i = a] dP_{A_i | \theta = \theta_1, L_e = l_e}(a) \\ &\geq \int_{\mathcal{Z}_+} P[T_i(Z_i) > t_u | \theta = \theta_1, A_i = a] dP_{A_i | \theta = \theta_1, L_e = l_e^-}(a). \end{aligned} \quad (7.50)$$

To prove (7.50) for any  $t_u$ , let  $h(a) := P[T_i(Z_i) > t_u | \theta = \theta_1, A_i = a]$ . Since  $h(a)$  is nonnegative, (7.50) is equivalent to

$$\int_0^\infty P_{A_i | \theta = \theta_1, L_e = l_e}(\{a : h(a) > y\}) dy \geq \int_0^\infty P_{A_i | \theta = \theta_1, L_e = l_e^-}(\{a : h(a) > y\}) dy, \quad (7.51)$$

which follows if,

$$\forall y \geq 0, P_{A_i|\theta=\theta_1, L_e=l_e}(\{a : h(a) > y\}) \geq P_{A_i|\theta=\theta_1, L_e=l_e^-}(\{a : h(a) > y\}). \quad (7.52)$$

It is now claimed that  $h(a)$  is a nondecreasing function of  $a$  on  $\mathcal{X}_+$ ; i.e.,  $\forall a^-, a^+ \in \mathcal{X}_+$  that satisfy  $a^- \leq a^+$ ,  $h(a^-) \leq h(a^+)$ . To see this, write

$$\begin{aligned} h(a^-) &= P[T_i(A_i + W_i) > t | \theta = \theta_1, A_i = a^-] = P[T_i(a^- + W_i) > t] \\ &\leq P[T_i(a^+ + W_i) > t] \end{aligned} \quad (7.53)$$

$$= P[T_i(A_i + W_i) > t | \theta = \theta_1, A_i = a^+] = h(a^+), \quad (7.54)$$

where (7.53) follows from condition (7.27) because both  $a^- + W_i, a^+ + W_i \in \mathcal{X}$  and  $a^- + W_i \leq a^+ + W_i$  for any realization of  $W_i$ .

Since  $h(a)$  is nondecreasing, for any  $y$ ,  $h(a) > y \Leftrightarrow a \in J_a$ , for either  $J_a = (a_y, \infty)$  or  $J_a = [a_y, \infty)$  and some  $a_y$ . Thus, (7.52) follows if

$$\forall a, P_{A_i|\theta=\theta_1, L_e=l_e}(J_a) \geq P_{A_i|\theta=\theta_1, L_e=l_e^-}(J_a), \quad (7.55)$$

and it is enough to prove (7.55) for  $J_a = (a, \infty)$  since the case  $J_a = [a, \infty)$  follows from Proposition F.3.

In order to prove (7.55), let  $X := \xi(\|L_i - L_e\|)$  and write

$$P[A_i \in J_a | \theta, L_e] = E[E[\mathbf{1}\{A_i \in J_a\} | \theta, L_e, L_i] | \theta, L_e] \quad (7.56)$$

$$= E[v(J_a | \theta \cdot X) | \theta, L_e], \quad (7.57)$$

where (7.56) follows from the Smoothing Property [54] and (7.57) follows from Assumption V of Chapter 5. Thus, for any  $l_e \in S_e$  and any  $\theta_1$ , write

$$P[A_i \in J_a | \theta = \theta_1, L_e = l_e] = \int v(J_a | \theta_1 \cdot x) dP_{X|L_e=l_e}(x) \quad (7.58)$$

and observe that  $P[X \leq x | L_e = l_e]$  forms a probability measure for all  $l_e \in S_e$ .

Therefore, (7.55) follows once it is proven that

$$\forall a, \int v(J_a | \theta_1 \cdot x) dP_{X|L_e=l_e}(x) \geq \int v(J_a | \theta_1 \cdot x) dP_{X|L_e=l_e^-}(x), \quad (7.59)$$

and since  $v(J_a | \theta_1 x) \geq 0$ , (7.59) follows once it is proven that

$$\forall a, \int_0^\infty P_{X|L_e=l_e}(\{x : v(J_a | \theta_1 \cdot x) > t\}) dt \geq \int_0^\infty P_{X|L_e=l_e^-}(\{x : v(J_a | \theta_1 \cdot x) > t\}) dt. \quad (7.60)$$

To prove (7.60), recall that it is enough to consider  $J_a = (a_y, \infty)$  and observe that  $v(J_a | \theta_1 x) := P[A_i > a_y | \theta = \theta_1, X = x]$ , which is assumed to be a nondecreasing function of  $x$  in this proposition. Thus,  $v(J_a | \theta_1 \cdot x) > t \Leftrightarrow x \in J_x$ , where  $J_x = [x_t, \infty)$  or  $J_x = (x_t, \infty)$  for some  $x_t$  that depends on  $t$  and  $a_y$ , and it is enough to prove the case  $J_x = (x_t, \infty)$



because the other case follows from Proposition F.3. Let  $B_r^\circ(l)$  represent the open ball of radius  $r$  centered at  $l$  and write

$$\begin{aligned} P_{X|L_e=l_e}(\{x : v(J_a|\theta_1 \cdot x) > t\}) &= P[X > x_t | L_e = l_e] \\ &= P[\xi(\|L_i - L_e\|) > x_t | L_e = l_e] \end{aligned} \quad (7.61)$$

$$= P[\xi(\|L_i - l_e\|) > x_t], \quad (7.62)$$

$$= P[\|L_i - l_e\| < d_x] = P[L_i \in B_{d_x}^\circ(l_e) \cap S_s], \quad (7.63)$$

where the first equality of (7.63) follows because  $\xi$  is assumed to be nondecreasing and right-continuous from Assumption VII of Chapter 5, which means that  $\xi(\|L_i - l_e\|) > x_t \Leftrightarrow \|L_i - l_e\| < d_x$  for some  $d_x$ .

Given (7.63), (7.60) follows if

$$\forall d, P[L_i \in B_d^\circ(l_e) \cap S_s] = \frac{m(B_d^\circ(l_e) \cap S_s)}{m(S_s)} \geq \frac{m(B_d^\circ(l_e^-) \cap S_s)}{m(S_s)} = P[L_i \in B_d^\circ(l_e^-) \cap S_s], \quad (7.64)$$

which follows because the area of the intersection of a convex region and a closed disk centered at  $l_e$  decreases as  $l_e$  moves away from the center, as proven in details in Lemma F.4 in the appendix.

To reach the third conclusion, replace  $l_e^-$  with  $l_e^+$  and reverse the order of the various inequalities between probabilities in all the steps except (7.53) in between (7.41) and (7.64).

To reach the second and third conclusions for a detection system belongs to the class  $\mathcal{D}_c$ , replace  $\{U_i\}_{i=1}^K$  with  $\{Z_i\}_{i=1}^K$  in all steps up to (7.43) and replace  $T_{0,i}(U_i, m)$  with  $T_{0,i}(Z_i)$  in the steps in between (7.44) and (7.45). The modified (7.45) is equivalent to (7.49), which was proven above. ■

Note that  $P[A_i > a_i | \theta = \theta_1, \xi(\text{dist}(L_i, L_e)) = x]$  is a nondecreasing function of  $x$  when  $A_i = \theta \xi(\text{dist}(L_i, L_e))$ , as in cooperative spectrum-sensing applications [94, 111, 120, 139], and when  $A_i$  is Poisson distributed with parameter given by  $\theta \xi(\text{dist}(L_i, L_e))$ , as in radiation detection applications [20, 58, 87, 98, 106], as shown in Proposition F.6 in the appendix.

Proposition 7.7 is one of the key results of this chapter. If a system designer adopts a detection system  $\underline{\phi}$  that belongs to the class of systems  $\mathcal{D}_c$  or  $\mathcal{D}_d$ , satisfies  $\alpha(\underline{\phi}) \leq \alpha_{\max}$ , and if the conditions of the proposition are met, then the system designer can determine a least favorable distribution  $P_{L_e}^-$  and a most favorable distribution  $P_{L_e}^+$  for  $\underline{\phi}$  and the performance under any distribution  $P_{L_e}$  is guaranteed to satisfy

$$\beta(P_{L_e}^+, \underline{\phi}) \geq \beta(P_{L_e}, \underline{\phi}) \geq \beta(P_{L_e}^-, \underline{\phi}). \quad (7.65)$$

Thus, if the system designer designs the detection system (i.e., determines the number of sensors  $K$ ) to achieve the prescribed detection performance while assuming that the emitter location has distribution  $P_{L_e}^-$ , then it can be ensured that the prescribed detection performance will be achieved under any distribution for the emitter location. Furthermore, the system designer can also determine how many sensors would be required to meet the prescribed detection performance while assuming  $L_e = (0, 0)$  in order to evaluate how conservative the resulting design is.

Proposition 7.7 contains as special cases the maximin and the maximax detection systems. Given its importance, this conclusion is presented using the following separate propositions.

**Proposition 7.8.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let  $Z_i$  assume values in  $\mathcal{Z} = \mathbb{R}$ ,  $\mathcal{Z} = \mathbb{R}_+$ ,  $\mathcal{Z} = \mathbb{Z}$ , or  $\mathcal{Z} = \mathbb{Z}_+$ ; and let  $\mathcal{Z}_+ := \mathcal{Z} \cap \mathbb{R}_+$ .<sup>7</sup> Assume  $f_W(w) > 0$  for any  $w \in \mathcal{Z}$  and consider that*

$$\forall w_1 \leq w_2 \in \mathcal{Z}, \forall a \in \mathcal{Z}_+, \frac{f_W(w_1 - a)}{f_W(w_1)} \leq \frac{f_W(w_2 - a)}{f_W(w_2)}; \quad (7.66)$$

*i.e.,  $f_W(w - a)/f_W(w)$  is nondecreasing in  $w \in \mathcal{Z}$  for all  $a \in \mathcal{Z}_+$ . Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , and define  $S_e^- := S_e \cap \partial B_{r_M}(0)$ . Assume a distribution  $P_{L_e^-}$  that satisfies  $P[L_e \in S_e^-] = 1$ .*

*Let  $\phi_{P_{L_e^-}}$  be the centralized detection system that maximizes the probability of detection for a given probability of false alarm when  $L_e$  is distributed according to  $P_{L_e^-}$ ; i.e.,  $\phi_{P_{L_e^-}}$  is given by (7.1).*

*Let  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$  be the distributed detection system that maximizes the probability of detection for a given probability of false alarm when  $L_e$  is distributed according to  $P_{L_e^-}$ ; i.e.,  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$  is given by (7.4) with  $\mathcal{C}_d$  being the class of all distributed detection systems.*

*If*

- 1. the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  for some  $c > 0$ ;*
- 2. for any  $a_i$ ,  $P[A_i > a_i | \theta = \theta_1, \xi(\mathbf{dist}(L_i, L_e)) = x]$  is a nondecreasing function of  $x$ ,*

*then*

- the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$ ;*
- a distribution for  $L_e$  such that  $P[L_e \in S_e^-] = 1$  is a least favorable distribution for either  $\phi_{P_{L_e^-}}$  or  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$ ; and*
- the detection system  $\phi_{P_{L_e^-}}$  is maximin centralized and  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$  is maximin for  $\mathcal{C}_d$ ; i.e., is maximin distributed.*

*Proof:* It will be shown that  $\phi_{P_{L_e^-}}$  and  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$  are elements of  $\mathcal{D}_c$  and  $\mathcal{D}_d$  respectively, and Proposition 7.7 can be used to reach the first and second conclusions. To see that  $\phi_{P_{L_e^-}} \in \mathcal{D}_c$ , observe that the optimum centralized detection system is a Neyman–Pearson detector, which has the likelihood ratio as the decision statistic, and  $\phi_{P_{L_e^-}}$  satisfies Proposition 7.4 with  $S_e' = S_e^-$ , which gives that  $\phi_{P_{L_e^-}} \in \mathcal{D}_c$ . To see that  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}} \in \mathcal{D}_d$ , observe that the distributed detection system  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$  is the optimum distributed detection system under  $P_{L_e^-}$  and, as shown in [114, 117, 123],

<sup>7</sup>As shown in the proof of Proposition 7.7, it follows that any realization of  $A_i$  belongs to  $\mathcal{Z}_+$  and any realization of  $W_i$  belongs to  $\mathcal{Z}$  for any of the specified sets  $\mathcal{Z}$ .

the optimal distributed detection system under the Neyman–Pearson framework when  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. is a distributed detection system with dependent randomization with fusion and sensor functions as defined in the Proposition 7.6 with  $S'_e = S_e^-$ ,<sup>8</sup> which gives that  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}} \in \mathcal{D}_d$ .

Now that the second conclusion is proven, it is possible to use the fact that the distribution that satisfies  $P[L_e \in S_e^-] = 1$  is a least favorable distribution for either  $\phi_{P_{L_e^-}}$  or  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$ . Since  $\phi_{P_{L_e^-}}$  and  $\phi_{\{P_{L_e^-}|\mathcal{C}_d\}}$  are optimal systems for  $\mathcal{C}_c$  and  $\mathcal{C}_d$  respectively, use Lemma 7.1 to prove the third conclusion. ■

An implicit conclusion of Proposition 7.8 is that, under the specified conditions, a least favorable distribution for the optimal centralized detection system is also a least favorable distribution for the optimal distributed detection system. The authors in [123] have also reached the same conclusion (in greater generality) when the measurements are conditionally independent. For the subset of cases that satisfy the specified conditions of Proposition 7.8, the results presented here extend the results of [123] to the case in which the composite hypothesis is formed by conditionally dependent distributions given by a mixture of i.i.d. distributions.

This conclusion is further extended in the next theorem, which has sufficient conditions to find the maximax detector and to conclude that a most favorable distribution for the optimal centralized detection system is also a most favorable distribution for the optimal distributed detection system.

**Proposition 7.9.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let  $Z_i$  assume values in  $\mathcal{Z} = \mathbb{R}$ ,  $\mathcal{Z} = \mathbb{R}_+$ ,  $\mathcal{Z} = \mathbb{Z}$ , or  $\mathcal{Z} = \mathbb{Z}_+$ ; and let  $\mathcal{Z}_+ := \mathcal{Z} \cap \mathbb{R}_+$ .<sup>9</sup> Assume  $f_W(w) > 0$  for any  $w \in \mathcal{Z}$  and consider that*

$$\forall w_1 \leq w_2 \in \mathcal{Z}, \forall a \in \mathcal{Z}_+, \frac{f_W(w_1 - a)}{f_W(w_1)} \leq \frac{f_W(w_2 - a)}{f_W(w_2)}, \quad (7.68)$$

*i.e.,  $f_W(w - a)/f_W(w)$  is nondecreasing in  $w \in \mathcal{Z}$  for all  $a \in \mathcal{Z}_+$ . Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , and define  $S_e^+ := \{0\}$ . Assume a distribution  $P_{L_e^+}$  that satisfies  $P[L_e \in S_e^+] = 1$ .*

*Let  $\phi_{P_{L_e^+}}$  be the centralized detection system that maximizes the probability of detection for a given probability of false alarm when  $L_e$  is distributed according to  $P_{L_e^+}$ ; i.e.,  $\phi_{P_{L_e^+}}$  is given by (7.1).*

<sup>8</sup>More precisely, the fusion and sensor functions defined in the Proposition 7.6 with  $S'_e = S_e^-$  satisfy Theorem 3.2 in Chapter 3. Although Theorem 3.2 gives the conditions for the optimal distributed detection system when  $\forall i, y, U_{i,y}^{\max} = U^{\max}$ , the system defined in Proposition 7.6 is equivalent to a system that satisfies  $\forall i, y, U_{i,y}^{\max} = U^{\max}$ . To see this, keep the fusion function and the dependent randomization as defined in Proposition 7.6 and change the sensor function as follows: pick  $U^{\max}$  as being the maximum among all  $U_{i,y}^{\max}$ , augment the existing intervals with intervals that satisfy  $\forall u > U_{i,y}^{\max}, I_{i,y,u} = \{\infty\}$ , change the sensor functions so that

$$\phi_i(z_i, y) = \min\{u \in \{0, \dots, U^{\max}\} : \mathcal{L}_{Z_i, S_e^-}(z_i) \in I_{i,y,u}\}, \quad (7.67)$$

and observe that  $P[U_i = u | \theta = \theta^*, Y = y] = 0$  for any  $u > U_{i,y}^{\max}$  and  $\theta^* \in \{0, \theta_1\}$  because the existing intervals are kept unchanged and their union equals  $\mathbb{R}_+ \cup \{\infty\}$ ; thus, the introduction of the intervals to satisfy  $\forall i, y, U_{i,y}^{\max} = U^{\max}$  does not change the system performance.

<sup>9</sup>As shown in the proof of Proposition 7.7, it follows that any realization of  $A_i$  belongs to  $\mathcal{Z}_+$  and any realization of  $W_i$  belongs to  $\mathcal{Z}$  for any of the specified sets  $\mathcal{Z}$ .

Let  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}_d\}}$  be the distributed detection system that maximizes the probability of detection for a given probability of false alarm when  $L_e$  is distributed according to  $P_{L_e}^+$ ; i.e.,  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}_d\}}$  is given by (7.4) with  $\mathcal{C}_d$  being the class of all distributed detection systems.

If

1. the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  for some  $c > 0$ ;
2. for any  $a_i$ ,  $P[A_i > a_i | \theta = \theta_1, \xi(\text{dist}(L_i, L_e)) = x]$  is a nondecreasing function of  $x$ ,

then

- the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$ ;
- a distribution for  $L_e$  such that  $P[L_e \in S_e^+] = 1$  is a most favorable distribution for either  $\underline{\phi}_{P_{L_e}^+}$  or  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}_d\}}$ ; and
- the detection system  $\underline{\phi}_{P_{L_e}^+}$  is maximax centralized and  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}_d\}}$  is maximax for  $\mathcal{C}_d$ ; i.e., is maximax distributed.

*Proof:* Using similar arguments as in the proof of Proposition 7.8,<sup>10</sup> one reaches that  $\underline{\phi}_{P_{L_e}^+}$  and  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}_d\}}$  respectively satisfy the conditions of Propositions 7.4 and 7.6 with  $S_e' = S_e^+$ . Such propositions indicate that  $\underline{\phi}_{P_{L_e}^+}$  and  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}_d\}}$  are elements of  $\mathcal{D}_c$  and  $\mathcal{D}_d$  respectively, which allows one to use Proposition 7.7 to reach the first and second conclusions.

To reach the third conclusion, Lemma 7.2 is used. The conditions on  $\mathcal{C}$  and  $\underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}$  in Lemma 7.2 are satisfied by construction for either  $\mathcal{C} = \mathcal{C}_c$  or  $\mathcal{C} = \mathcal{C}_d$ . It remains to show that, for both  $\mathcal{C}_c$  and  $\mathcal{C}_d$ , condition (7.16) of Lemma 7.2 is satisfied. To show this, pick any distribution  $P_{L_e}'$  and the detection system  $\underline{\phi}_{\{P_{L_e}'|\mathcal{C}\}}$  that satisfies (7.4) for either  $\mathcal{C} = \mathcal{C}_c$  or  $\mathcal{C} = \mathcal{C}_d$ ; i.e.,  $\underline{\phi}_{\{P_{L_e}'|\mathcal{C}\}}$  is either the optimum centralized detection system or optimum distributed detection system under  $P_{L_e}'$ . Let  $\underline{\phi}_{\{P_{L_e}^{(l_e)}|\mathcal{C}\}}$  denote the system that satisfies (7.4) when  $L_e$  has the distribution  $P_{L_e}^{(l_e)}$  that satisfies  $P[L_e = l_e] = 1$ .

It is claimed that  $P_{L_e}^+$  is a most favorable distribution for both  $\underline{\phi}_{\{P_{L_e}^{(l_e)}|\mathcal{C}_c\}}$  and  $\underline{\phi}_{\{P_{L_e}^{(l_e)}|\mathcal{C}_d\}}$ . To see this, recall that optimum centralized or distributed detection systems follow the form of detection systems in Propositions 7.4 or 7.6, which means that  $\underline{\phi}_{\{P_{L_e}^{(l_e)}|\mathcal{C}\}} \in \mathcal{D}$ , and use Proposition 7.7.

With these definitions, one can write

$$\beta(P_{L_e}', \underline{\phi}_{\{P_{L_e}'|\mathcal{C}\}}) = \int \beta(P_{L_e}^{(l_e)}, \underline{\phi}_{\{P_{L_e}'|\mathcal{C}\}}) dP_{L_e}'(l_e) \leq \int \beta(P_{L_e}^{(l_e)}, \underline{\phi}_{\{P_{L_e}^{(l_e)}|\mathcal{C}\}}) dP_{L_e}'(l_e) \quad (7.69)$$

$$\leq \int \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^{(l_e)}|\mathcal{C}\}}) dP_{L_e}'(l_e) \leq \int \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}) dP_{L_e}'(l_e) = \beta(P_{L_e}^+, \underline{\phi}_{\{P_{L_e}^+|\mathcal{C}\}}), \quad (7.70)$$

<sup>10</sup>Just replace  $P_{L_e}^-$  with  $P_{L_e}^+$  and replace  $S_e^-$  with  $S_e^+$ .

where the inequality in (7.69) follows because  $\underline{\phi}_{\{P_{L_e}^{(l_e)}\}|\mathcal{C}}$  is the optimal detector in  $\mathcal{C}$  under  $P_{L_e}^{(l_e)}$ , the first inequality in (7.70) follows because  $P_{L_e}^+$  is a most favorable distribution for  $\underline{\phi}_{\{P_{L_e}^{(l_e)}\}|\mathcal{C}}$ , and the second inequality in (7.70) follows because  $\underline{\phi}_{\{P_{L_e}^+\}|\mathcal{C}}$  is the optimal detector in  $\mathcal{C}$  under  $P_{L_e}^+$ . ■

The following observations can be made:

- Comparing the conditions of Propositions 7.8 and 7.9, it is possible to note a common set of conditions that make  $P[L_e \in S_e^-] = 1$  define a least favorable distribution for maximin systems and make  $P[L_e \in S_e^+] = 1$  define a most favorable distribution for maximax systems;
- Propositions 7.7, and 7.8 allow one to find the detection systems that are either maximin or maximin for the class of distributed detection systems when the family of distributions is different from the family generally considered in literature, which is composed of a nominal distribution and distributions in its neighborhood [59].
- $f_W(w - a)/f_W(w)$  is monotonically nondecreasing in  $w \in \mathcal{Z}$  for all  $a \in \mathcal{Z}_+$  if  $W$  is a Poisson distribution, a Laplace distribution, a Gaussian distribution, or any other Generalized Gaussian distribution with parameter at least 1.

Some may consider that the conclusions of Proposition 7.8 are trivial since the distribution that places the emitter location in the boundary of  $S_e$  with probability one; i.e.,  $P[L_e \in S_e^-] = 1$ ; maximizes the expected distance between any sensor and the emitter.

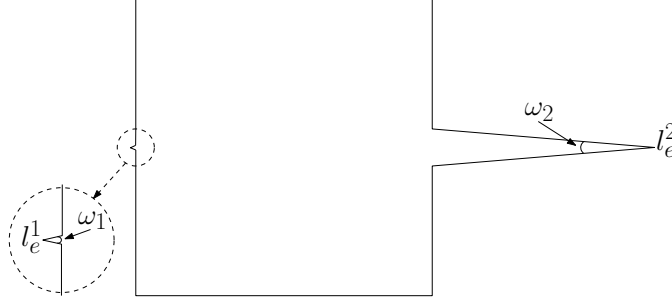
Furthermore, some may consider that the distribution of the noise random variables should not influence a least or a most favorable distribution and question the need of the assumption on  $f_W(w)$  in Propositions 7.8 and 7.9.

The next two subsections contain two examples that illustrate that the conclusions of Proposition 7.7 are not trivial and the condition of  $f_W(w)$  is not superfluous.

### 7.3.1 Example 7.1:

This subsection contains an example that illustrates that the distribution for the emitter location that maximizes the expected distance between any sensor and the emitter is not always a least favorable distribution for a detection system.

Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let  $A_i = \theta_1 \xi(\|L_i - L_e\|)$  with  $\xi(r) = \mathbf{1}\{r < r_0\}$ , where  $r_0 = 0.1 - \varepsilon$  for any  $\varepsilon \in (0, 0.1)$ , and let  $\theta_1 = 1$ . Let  $F_W$  be increasing. Assume that  $S_e$  is the union of a square  $S_q$  and two triangles isosceles  $S_t^1$  and  $S_t^2$ . Let  $S_q$  be centered at the origin and have sides of length 2. Let  $S_t^1$  and  $S_t^2$  have the line  $y = 0$  as their line of symmetry. Let  $S_t^1$  have its base located at the line  $x = -1$  and have a vertex at  $l_e^1 = (-1.1, 0)$  with an angle  $\omega_1 = \pi/100$ . Let  $S_t^2$  have its base located at the line  $x = 1$  and have a vertex at  $l_e^2 = (1 + \sqrt{8}, 0)$  with an angle  $\omega_2 = 2\omega_1$ . Figure 7.1 illustrates the shape of  $S_e$ . Observe that such a region of interest does not satisfy the conditions of Propositions 7.7 or 7.8. Assume



**Figure 7.1.** Region of interest  $S_e$  given by the union of a square and two triangles in which the distribution for the emitter location that places the emitter at  $l_e^2$  is not always a least favorable distribution for a detection system.

that  $S_s = S_e$ , consider a single sensor ( $K = 1$ ), and assume that the sensor location  $L_1$  is uniformly distributed in  $S_s$ . Consider a centralized detection system in which the fusion function is given by  $\phi_0(z_1, y_0) = \mathbf{1}\{z_1 > t_0\}$ . Let  $t_0$  be such that  $P[Z_1 > t_0 | \theta = 0] = \alpha_{\max}$  for some  $\alpha_{\max} \in (0, 1)$ .

Consider that the emitter location  $L_e$  has a distribution  $P_{L_e}^{(l_e^2)}$  that satisfies  $P[L_e = l_e^2] = 1$ . It may be considered intuitive that such a distribution would provide a least favorable distribution for a given detection system because  $l_e^2$  is the most distant point from the origin in  $S_e$ . It will be shown that, under the conditions of this example,  $P_{L_e}^{(l_e^2)}$  is *not* a least favorable distribution for  $\phi_0$ .

From the definition of a least favorable distribution for a detection system, (7.6),  $P_{L_e}^{(l_e^2)}$  is a least favorable distribution for  $\phi_0$  if and only if

$$\forall P_{L_e}, \beta(P_{L_e}, \phi_0) \geq \beta(P_{L_e}^{(l_e^2)}, \phi_0), \quad (7.71)$$

which follows if

$$\forall l_e \in S_e, \beta(P_{L_e}^{(l_e)}, \phi_0) \geq \beta(P_{L_e}^{(l_e^2)}, \phi_0), \quad (7.72)$$

since  $\beta(P_{L_e}, \phi_0) = \int \beta(P_{L_e}^{(l_e)}, \phi_0) dP_{L_e}(l_e)$ .

It is possible to show that

$$\begin{aligned} \beta(P_{L_e}^{(l_e)}, \phi_0) &= P[L_1 \in (S_s \cap B_{r_0}^\circ(l_e))] \cdot [F_W(t_0) - F_W(t_0 - 1)] + 1 - F_W(t_0) \\ &= \frac{m(S_s \cap B_{r_0}^\circ(l_e))}{m(S_s)} \cdot [F_W(t_0) - F_W(t_0 - 1)] + 1 - F_W(t_0), \end{aligned} \quad (7.73)$$

where  $B_r^\circ(l)$  represents the open ball with radius  $r$  and centered at  $l$ , as shown in details in Appendix G.

Since  $F_W(t_0) - F_W(t_0 - 1) > 0$ ,

$$\forall l_e', l_e'' \in S_e, \beta(P_{L_e}^{(l_e')}, \phi_0) \geq \beta(P_{L_e}^{(l_e'')}, \phi_0) \Leftrightarrow m(S_s \cap B_{r_0}^\circ(l_e')) \geq m(S_s \cap B_{r_0}^\circ(l_e'')), \quad (7.74)$$

and (7.72) is equivalent to

$$\forall l_e \in S_e, m(S_s \cap B_{r_0}^\circ(l_e)) \geq m(S_s \cap B_{r_0}^\circ(l_e^2)). \quad (7.75)$$

The distribution  $P_{L_e}^{(l_e^2)}$  is not a least favorable distribution for  $\phi_0$  because  $m(S_s \cap B_{r_0}^\circ(l_e^1)) < m(S_s \cap B_{r_0}^\circ(l_e^2))$ . To see this, recall that  $r_0 = 0.1 - \varepsilon$  for some  $\varepsilon \in (0, 0.1)$ , which means that the intersection between  $B_{r_0}^\circ(l_e^1)$  and the square  $S_q$  is empty. For the same reason, the intersection between  $B_{r_0}^\circ(l_e^2)$  and the square  $S_q$  is also empty. This means that  $S_s \cap B_{r_0}^\circ(l_e^1)$  and  $S_s \cap B_{r_0}^\circ(l_e^2)$  are sectors of radius  $r_0$  and angles  $\omega_1$  and  $\omega_2$ , respectively. Since  $\omega_1 < \omega_2$ ,  $m(S_s \cap B_{r_0}^\circ(l_e^1)) < m(S_s \cap B_{r_0}^\circ(l_e^2))$ .<sup>11</sup>

This example highlights that a distribution for the emitter location that is a least favorable distribution for a detection system cannot be guessed by simply observing the shape of the region of interest.

### 7.3.2 Example 7.2:

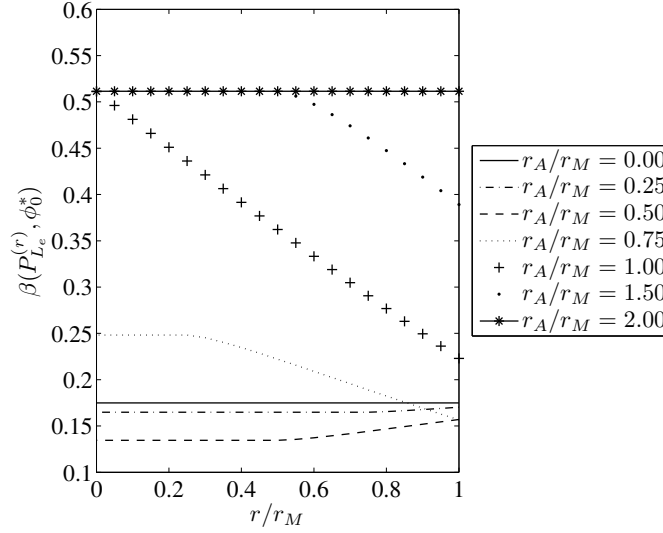
Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Consider a single sensor, let  $\theta_1 = 1$ , and take  $A_1 = \xi(\|L_e - L_1\|)$ , with  $\xi(r) = 4$  for  $r \in [0, r_A)$ ,  $\xi(r) = 2$  for  $r \in [r_A, 2r_M)$  and  $\xi(r) = 0$  for  $r \geq 2r_M$ . Let  $W$  be distributed according to the Cauchy distribution; i.e.,  $f_W(w) = (\lambda/\pi)/(\lambda^2 + w^2)$ , and assume  $\lambda = 0.5$ . Let the region of interest  $S_e = B_{r_M}(0)$  and assume that  $L_1$  is uniformly distributed on  $S_s = S_e$ . Suppose that a system designer adopts the distribution  $P_{L_e}^*$  that satisfies  $P[L_e = (0, r_M)] = 1$  and builds a centralized detection system with fusion function  $\phi_0^*(\mathbf{z}, y)$  given by (7.23), in which  $S_e' = \{(0, r_M)\}$ . Note that  $\phi_0^*$  is the optimum detection system under the distribution  $P_{L_e}^*$ ; i.e., the resulting sensor detection system equals the system  $\phi_{P_{L_e}^*}$  given by (7.1). The threshold  $t_{0,1}$  of  $\phi_0^*$  is chosen so that the resulting probability of false alarm equals 0.01, where it is noted that no randomization is necessary to achieve this probability of false alarm.

With such assumptions, the only condition of Proposition 7.8 that is not satisfied is relation (7.66); i.e.,  $f_W(w - a)/f_W(w)$  is not nondecreasing in  $w \in \mathcal{Z}$  for all  $a \in \mathcal{Z}_+$  because  $W$  has the Cauchy distribution.

It will be shown that, although  $P_{L_e}^*$  satisfies  $P[L_e \in S_e^-] = 1$ , it is not a least favorable distribution for  $\phi_0^*$  in some cases. To show this, the probability of detection achieved by  $\phi_0^*$  will be obtained under other distributions for  $L_e$ .

Using numerical integration, it is possible to obtain Figure 7.2, which shows the probability of detection achieved by  $\phi_0^*$  when  $L_e$  is distributed according to  $P_{L_e}^{(r)}$  for various  $r \leq r_M$  under several ratios of  $r_A/r_M$ , where  $P_{L_e}^{(r)}$  is the distribution that satisfies  $P[L_e = (0, r)] = 1$ . The ratio  $r_A/r_M$  defines the form of the function  $\xi$  and, therefore, influences the conditional distribution  $P_{Z_1|\theta=1, L_e=(0,r)}$  for any  $r$ . Note also that the ratio  $r_A/r_M$  influences the form of  $\phi_0^*$ .

<sup>11</sup>Note that  $P_{L_e}^{(l_e^2)}$  may be a least favorable distribution for  $\phi_0$  under other conditions. For instance, if  $r_0 = 2$  and  $\omega_2 = \pi/50$ , then  $P_{L_e}^{(l_e^2)}$  is a least favorable distribution for  $\phi_0$  since  $\forall l_e \in S_e, m(S_s \cap B_{r_0}^\circ(l_e)) \geq m(S_s \cap B_{r_0}^\circ(l_e^2)) = \omega_2 r_0^2/2 \approx 0.1257$ . To prove this claim, first observe that  $m(S_s \cap B_{r_0}^\circ(l_e)) \geq \pi$  for all  $l_e \in S_q$ . For the  $l_e \in S_e - S_q$ , observe that, for all  $l_e \in S_t^1$ ,  $m(S_s \cap B_{r_0}^\circ(l_e)) \geq m(S_s \cap B_{r_0}^\circ(l_e^1))$ . To see this, note that the sector with radius  $r_0$  and angle formed from  $-\omega_1/2$  and  $+\omega_1/2$  is always contained in  $S_e$  as one varies the center of the sector within  $S_t^1$ . When centered at  $l_e^1$ , such a sector is equal to  $S_s \cap B_{r_0}^\circ(l_e^1)$ ; and, when centered at any other  $l_e \in S_t^1$ , such a sector is a subset of  $S_s \cap B_{r_0}^\circ(l_e)$ . Likewise, the sector with radius  $r_0$  and angle formed from  $\pi - \omega_2/2$  and  $\pi + \omega_2/2$  is always contained in  $S_e$  as one varies the center of the sector within  $S_t^2$ , which means that, for all  $l_e \in S_t^2$ ,  $m(S_s \cap B_{r_0}^\circ(l_e)) \geq m(S_s \cap B_{r_0}^\circ(l_e^2))$ . Thus, it suffices to show that  $m(S_s \cap B_{r_0}^\circ(l_e^1)) \geq m(S_s \cap B_{r_0}^\circ(l_e^2))$ . Observe that  $S_s \cap B_{r_0}^\circ(l_e^1)$  contains the triangle formed by the points  $(-1, +\sqrt{r_0^2 - 0.1^2})$ ,  $(-1, -\sqrt{r_0^2 - 0.1^2})$ , and  $(-1.1 + r_0, 0)$ , and the area of such a triangle equals  $(r_0 - 0.1)\sqrt{r_0^2 - 0.1^2}$ . Thus, since  $r_0 = 2$ ,  $m(S_s \cap B_{r_0}^\circ(l_e^1)) > (r_0 - 0.1)\sqrt{r_0^2 - 0.1^2} \approx 3.8 > m(S_s \cap B_{r_0}^\circ(l_e^2))$ .



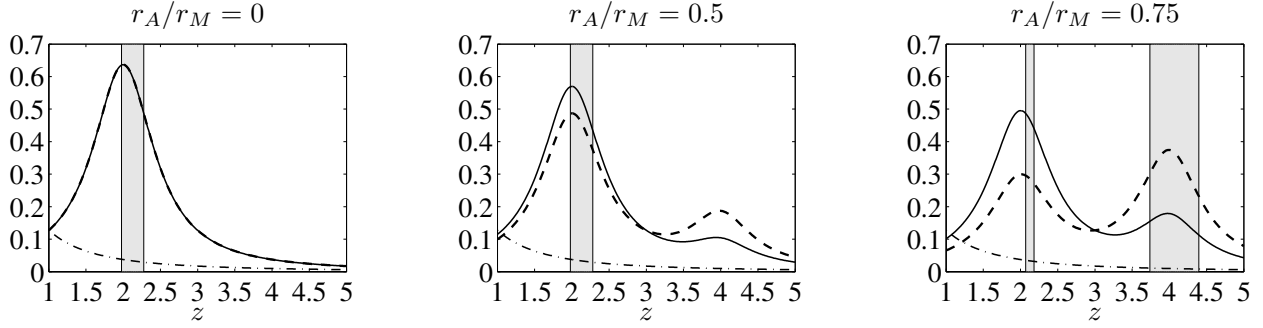
**Figure 7.2.** Probability of detection under 0.01 probability of false alarm achieved by the centralized detection system  $\phi_0^*$  under the distribution  $P_{L_e}^{(r)}$ , where  $P_{L_e}^{(r)}$  satisfies  $P[L_e = (0, r)] = 1$ , for various values  $r \leq r_M$  and for various ratios  $r_A/r_M$ , where  $r_A$  defines the function  $\xi$  in in Example 7.2.

From Figure 7.2, it is possible to see that, if  $r_A/r_M = 0.5$  or  $r_A/r_M = 0.25$ , then  $P_{L_e}^*$  is *not* a least favorable distribution for  $\phi_0^*$  because its probability of detection under  $P_{L_e}^*$ , which corresponds to the case  $r/r_M = 1$  is greater than the probability of detection under  $P_{L_e}^{(r)}$ , for any  $r/r_M < 1$ . In other words,  $P_{L_e}^*$  does not satisfy (7.6).

Figure 7.3 helps to explain why  $P_{L_e}^*$  is not a least favorable distribution for  $\phi_0^*$  for some of the cases. This figure presents the conditional p.d.f.s  $f_{Z_1|\theta=0}(z)$ ,  $f_{Z_1|\theta=1, L_e=(0, r_M)}(z)$ , and  $f_{Z_1|\theta=1, L_e=(0, 0)}(z)$ ; and the shaded regions indicate the  $H_1$  decision intervals of  $\phi_0^*$  for three cases of  $r_A/r_M$ . Focusing on the middle graph, which corresponds to the case  $r_A/r_M = 0.5$ , it is possible to observe that, when  $f_{Z_1|\theta=1, L_e=(0, 0)}(z)$  is integrated over the decision interval  $H_1$ , the resulting probability will be lower than the probability that results when integrating  $f_{Z_1|\theta=1, L_e=(0, r_M)}(z)$  over the same interval. This rationale also helps to explain the interesting phenomenon in that the probability of detection for the case  $r_A/r_M = 0.5$  is lower than the probability of detection for the  $r_A/r_M = 0$  case even though the case  $r_A/r_M = 0.5$  offers higher values for  $\xi$ . By comparing the curves corresponding to the  $r_A/r_M = 0.5$  case and the curves corresponding to the  $r_A/r_M = 0$  case, it is possible to observe that the density function under  $H_1$  for the  $r_A/r_M = 0$  case is higher than the density functions under  $H_1$  for the  $r_A/r_M = 0.5$  case when  $z$  belongs to the  $H_1$  decision interval.

In other words,  $P_{L_e}^*$  is not a least favorable distribution for  $\phi_0^*$  because the  $H_1$  decision region obtained under  $P_{L_e}^*$  does not cover the region around 4, which has increased probability mass under  $P_{L_e}^{(r)}$  as  $r$  decreases. This phenomenon disappears as  $r_A/r_M$  increases because the  $H_1$  decision region eventually becomes the union of two disjoint intervals, one close to 2 and another around 4, as shown in the right graph of Figure 7.3 for the case  $r_A/r_M = 0.75$ ; and, as  $r_A/r_M$  approaches 1, it eventually becomes a single interval around 4.





**Figure 7.3.** Conditional probability density functions  $f_{Z_1|\theta=0}(z)$  (dashed and dotted line),  $f_{Z_1|\theta=1,R=r_M}(z)$  (full line),  $f_{Z_1|\theta=1,R=0}(z)$  (dashed line), and  $H_1$  decision interval (shaded region) for 0.01 probability of false alarm when  $r_A/r_M = 0$  (left),  $r_A/r_M = 0.5$  (center), and  $r_A/r_M = 0.75$  (right) in Example 7.2.

Similar observations can be reached when considering most favorable distributions. Consider now that a system designer adopts the distribution  $P_{L_e}^{**}$  that satisfies  $P[L_e = (0, 0)] = 1$  to obtain an upper bound for the performance of any system. Let  $\phi_0^{**}$  be given by (7.23), in which  $S'_e = \{(0, 0)\}$ , with  $t_{0,1}$  configured such that the resulting probability of false alarm equals 0.01. The only condition of Proposition 7.9 that is not satisfied is relation (7.68). Although  $P_{L_e}^{**}$  satisfies  $P[L_e \in S_e^+] = 1$ , it is not a most favorable distribution for  $\phi_0^{**}$  when  $r_A/r_M = 0.25$ , as shown in Figure 7.4.

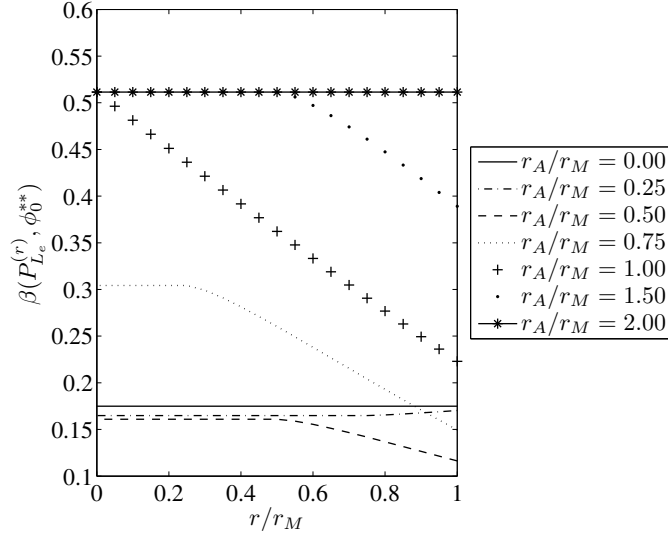
## 7.4 Summary

In this section, the concept of a least favorable distribution was applied to the emitter location distribution.

It was shown that the concept of a least favorable distribution usually considered in the literature was not sufficient for the design of sensor detection systems for two reasons: it does not allow the consideration of distributed detection systems; and it does not reflect the design objective of adopting a conservative design. The concepts of a least favorable distribution for a class of systems was proposed to address the first issue, and the concept of a least favorable distribution for a system was proposed to address both issues.

The concept of a most favorable distribution for the emitter location was defined. A most favorable distribution for a detection system can be used to find an upper bound for the detection performance of a system under any distribution; thus, when using both a least and a most favorable distribution, a system designer can determine a range for the detection performance of a system under any emitter location distribution.

Sufficient conditions were identified under which a least and a most favorable distributions for a detection system have simple forms in both practical and optimal detection systems. Under these conditions, a least favorable distribution for a detection system places the emitter on a subset of the boundary of the region of interest; and a most favorable distribution for a detection system places the emitter at the center of the region of interest. Intuition might suggest that these results are obvious since a distribution that places the emitter on the boundary of the region of interest maximizes



**Figure 7.4.** Probability of detection under 0.01 probability of false alarm achieved by the optimum centralized detection system  $\phi_0^{**}$  under the distribution  $P_{L_e}^{(r)}$ , where  $P_{L_e}^{(r)}$  satisfies  $P[L_e = (0, r)] = 1$ , for various values  $r \leq r_M$  and for various ratios  $r_A/r_M$ , where  $r_A$  defines the function  $\xi$  in Example 7.2.

the expected distance between sensors and emitter, suggesting that such a distribution is least favorable; and the distribution that places the emitter at the center of the region of interest minimizes the expected distance between sensors and emitter, suggesting that such a distribution is most favorable; however, as shown by examples, these distributional forms may not be least or most favorable for the detection system.

The results of this section also uncover an important additional benefit of adopting least or most favorable distributions for the emitter location: the measurements become conditionally i.i.d. under either a least or a most favorable distribution. This is perhaps the most important contribution of this chapter given that conditional dependence is a common obstacle for the design of a detection system.

## Chapter 8

### Using a Most Favorable Emitter Location Distribution to Evaluate How Conservative the Design Based on a Least Favorable Emitter Location Distribution is

As mentioned in Section 7.2, a most favorable distribution for a detection system  $\underline{\phi}$  can be used along with a least favorable distribution for  $\underline{\phi}$  to determine a range for the detection performance of a system as the emitter location distribution changes because

$$\forall P_{L_e}, \beta(P_{L_e}^+, \underline{\phi}) \geq \beta(P_{L_e}, \underline{\phi}) \geq \beta(P_{L_e}^-, \underline{\phi}), \quad (8.1)$$

where the reader should recall that  $\beta(P_{L_e}, \underline{\phi})$  represents the probability of detection obtained by the detection system  $\underline{\phi}$  under the assumption that the emitter location  $L_e$  is distributed according to  $P_{L_e}$ ; and  $P_{L_e}^-$  and  $P_{L_e}^+$  respectively represent a least and a most favorable distributions for the detection system  $\underline{\phi}$  being considered.

While relation (8.1) provides lower and upper bounds for the detection performance when considering a fixed number of sensors  $K$ , it is possible to consider a fixed detection performance and use least and most favorable distributions to determine lower and upper bounds for the required number of sensors to achieve the specified detection performance.

Since each possible value for  $K$  defines a different detection system, it will be useful to think of the system designer as considering a *sequence of detection systems*  $\Phi := \{\underline{\phi}^K\}_{K=1}^\infty := \{(\{\phi_{i,K}, P_{Y_{i,K}}\}_{i=0}^K)\}_{K=1}^\infty$ .

Suppose that the system designer can determine a pair of distributions  $P_{L_e}^-$  and  $P_{L_e}^+$  that are a least and a most favorable distribution for each detection system of the sequence  $\Phi$ ; i.e., for any  $K$ , a common pair of  $P_{L_e}^-$  and  $P_{L_e}^+$  forms a least and a most favorable distribution for the detection system  $\underline{\phi}^K$ . At least for the models considered in this dissertation, this assumption is not strong because the conditions of Proposition 7.7 are satisfied for all  $K$  if each detection system in the sequence  $\Phi$  belongs to either the class  $\mathcal{D}_c$  or the class  $\mathcal{D}_d$  of detection systems defined in Section 7.3. More precisely, for any  $K$ , if  $\underline{\phi}^K$  is a centralized detection system and each of the  $K$  functions  $\{T_{0,i}(z)\}_{i=1}^K$  that define the fusion function through (7.19) is a nonnegative and nondecreasing function, then each detection system in  $\Phi$  will belong to  $\mathcal{D}_c$ . Likewise, for any  $K$ , if  $\underline{\phi}^K$  is a distributed detection system with dependent randomization and

the  $K$  sensor functions  $\phi_{i,K}(z_i, y)$  are nondecreasing real functions of  $z_i$ , and the functions  $T_{0,i}(u_i, y)$  that define each fusion function  $\phi_{0,K}(\mathbf{u}, y)$  are nonnegative and nondecreasing, then each detection system in  $\Phi$  will belong to  $\mathcal{D}_d$ . To illustrate, suppose that  $\Phi$  is a sequence of centralized detection systems in which, for each  $K$ , the detection system  $\underline{\phi}^K$  is of the form defined in Proposition 7.3 with different decision thresholds; i.e.,  $\phi_{0,K}(\mathbf{z}, y_0) = \mathbf{1}\{\sum_{i=1}^K z_i > t_{0,K}\}$ . If the conditions of Proposition 7.7 are met, then the distributions  $P_{L_e}^-$  and  $P_{L_e}^+$  identified in Proposition 7.7 are a least and a most favorable distributions for  $\underline{\phi}^K$  for all  $K$ .

Given a sequence of detection systems  $\Phi$ , any distribution  $P_{L_e}$ , a minimum probability of detection  $\beta_{\min} \in (0, 1)$  requirement, and a maximum probability of false alarm  $\alpha_{\max} \in (0, 1)$  requirement, define

$$K^{(\Phi, P_{L_e})} := \inf \{K : \beta(P_{L_e}, \underline{\phi}^K) \geq \beta_{\min}, \alpha(\underline{\phi}^K) \leq \alpha_{\max}\}, \quad (8.2)$$

where the convention that  $\inf \emptyset = \infty$  is adopted here.

If a sequence of detection systems  $\Phi$  satisfies  $K^{(\Phi, P_{L_e}^-)} < \infty$ , then

$$\forall P'_{L_e}, 1 \leq \frac{K^{(\Phi, P_{L_e}^-)}}{K^{(\Phi, P'_{L_e})}} \leq \frac{K^{(\Phi, P_{L_e}^-)}}{K^{(\Phi, P_{L_e}^+)}} \quad (8.3)$$

which follows from (8.1) as shown in Appendix I.1.

Similarly, it is possible to write relations that involve a sequence of maximin and maximax detection systems for a particular class of systems  $\mathcal{C}$  that satisfies Condition 7.1. For that, define  $\Phi_{P_{L_e}|\mathcal{C}} := \{\underline{\phi}_{P_{L_e}|\mathcal{C}}^K\}_{K=1}^\infty$ , where  $\underline{\phi}_{P_{L_e}|\mathcal{C}}^K$  is the detection system with  $K$  sensors that satisfies (7.4).

If  $P_{L_e}^-$  is a least favorable distribution for  $\underline{\phi}_{P_{L_e}|\mathcal{C}}^K$  for every  $K$ ; i.e., the pair formed by each detection system  $\underline{\phi}_{P_{L_e}|\mathcal{C}}^K$  and  $P_{L_e}^-$  satisfies the conditions of Lemma 7.1; and, for every  $K$ , if  $\underline{\phi}_{P_{L_e}|\mathcal{C}}^K$  and  $P_{L_e}^+$  satisfy the conditions of Lemma 7.2; i.e.,  $\underline{\phi}_{P_{L_e}|\mathcal{C}}^K$  is maximin for  $\mathcal{C}$  and  $\underline{\phi}_{P_{L_e}|\mathcal{C}}^K$  is maximax for  $\mathcal{C}$ ; and if  $K^{(\Phi_{P_{L_e}|\mathcal{C}}, P_{L_e}^-)} < \infty$ , then

$$\forall P'_{L_e}, 1 \leq \frac{K^{(\Phi_{P_{L_e}|\mathcal{C}}, P_{L_e}^-)}}{K^{(\Phi_{P_{L_e}|\mathcal{C}}, P'_{L_e})}} \leq \frac{K^{(\Phi_{P_{L_e}|\mathcal{C}}, P_{L_e}^-)}}{K^{(\Phi_{P_{L_e}|\mathcal{C}}, P_{L_e}^+)}} \quad (8.4)$$

as shown in Appendix I.1.

Whether the system designer considers a particular sequence of detectors  $\Phi$  or a sequence of detectors  $\Phi_{P_{L_e}|\mathcal{C}}$  that is optimal for a given  $P'_{L_e}$ , it is difficult to determine  $K^{(\Phi, P_{L_e}^-)}/K^{(\Phi, P'_{L_e})}$  or  $K^{(\Phi_{P_{L_e}|\mathcal{C}}, P_{L_e}^-)}/K^{(\Phi_{P_{L_e}|\mathcal{C}}, P'_{L_e})}$  for all possible  $P'_{L_e}$ , particularly because the measurements may not be conditionally i.i.d. under  $P'_{L_e}$ ; therefore, the evaluation of how conservative the least favorable distribution is will be done through the looser upper bounds of (8.3) and (8.4). The benefit of focusing on the looser upper bounds of (8.3) and (8.4) is that the measurements may be conditionally i.i.d. under either  $P_{L_e}^-$  or  $P_{L_e}^+$ .

If the system designer is able to determine the ratio  $K^{(\Phi, P_{L_e}^-)}/K^{(\Phi, P_{L_e}^+)}$ , then the system designer can be assured that a design based on a least favorable distribution for the  $\Phi$  will not require more than  $K^{(\Phi, P_{L_e}^-)}/K^{(\Phi, P_{L_e}^+)}$  times the

number of sensors that would be deployed if the actual distribution for  $L_e$  were known. A similar statement can be made regarding the ratio  $K^{(\Phi_{P_{L_e}^-}|\mathcal{G}, P_{L_e}^-)} / K^{(\Phi_{P_{L_e}^+}|\mathcal{G}, P_{L_e}^+)}$ .

It is difficult to analytically determine the ratios  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^+)}$  and  $K^{(\Phi_{P_{L_e}^-}|\mathcal{G}, P_{L_e}^-)} / K^{(\Phi_{P_{L_e}^+}|\mathcal{G}, P_{L_e}^+)}$ ; and numerical techniques are usually required. In the next section, an example is presented in which the ratio  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^+)}$  is obtained numerically.

## 8.1 Example

Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let the region of interest  $S_e$  be a disk centered at the origin with radius  $r_M$ ; i.e.,  $S_e = B_{r_M}(0)$ . Let  $S_s = cS_e$  for  $c = 1$ , and let  $\{L_i\}_{i=1}^K$  be i.i.d. with the uniform distribution in  $S_s$ . For any  $l_i$  and  $l_e$ , let  $A_i$  conditioned on  $L_i = l_i$  and  $L_e = l_e$  be Poisson-distributed with parameter  $\theta_1 \xi(\|l_i - l_e\|)$  with  $\xi(r) = A_{\max}/(1+r)^2$  and  $\theta_1 = 1$ . Let  $W_i$  be Poisson-distributed with parameter  $\lambda_W > 0$ . For each  $K$ , let  $\underline{\phi}^K$  be the binary distributed detection system with given binary sensor functions

$$\phi_{i,K}(z_i, y) = \mathbf{1}\{z_i > t_s\}, \quad (8.5)$$

assume that  $K$  dedicated and error-free communication channels are available; and let the fusion function be

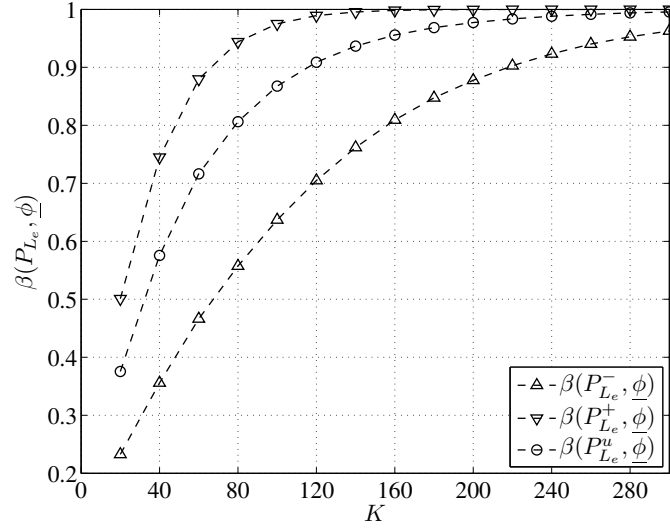
$$\phi_{0,K}(\mathbf{u}, y_0) = \mathbf{1}\left\{\sum_{i=1}^K u_i \in I_{0,y_0}\right\}, \quad (8.6)$$

where  $I_{0,0} = (k, \infty]$ ,  $I_{0,1} = [k, \infty]$ , and  $y_0$  is the realization of the randomization random variable  $Y_0 \in \{0, 1\}$ . From Theorem 3.1, (8.6) defines the optimal fusion function for the set of sensor functions (8.5) when  $U_i$  are conditionally i.i.d. given  $\theta$ .

Consider any distribution  $P_{L_e}^-$  that satisfies  $P[L_e \in S_e^*] = 1$  for  $S_e^* = \partial B_{r_M}(0)$ ; i.e.,  $P_{L_e}^-$  places the emitter in the boundary of the disk with probability 1. From Proposition 7.7,  $P_{L_e}^-$  is a least favorable emitter location distribution for  $\underline{\phi}^K$  for any  $K$ . From the same proposition, it also follows that the distribution  $P_{L_e}^+$  that satisfies  $P[L_e = (0, 0)] = 1$  is a most favorable emitter location distribution for  $\underline{\phi}^K$  for any  $K$ .

As shown in Appendix H, it is possible to use numerical integration to evaluate the probabilities of detection under  $P_{L_e}^-$ , under  $P_{L_e}^+$ , and under the uniform distribution  $P_{L_e}^u$  for various values of  $K$ . Taking  $\alpha_{\max} = 0.05$ ,  $\lambda_W = 1$ ,  $t_s = \lambda_W$ ,  $r_M = 15$ , and  $A_{\max} = 40$ , the resulting probabilities of detection are shown in Figure 8.1.

From Figure 8.1, it is possible to see that, if a system designer imposes a minimum probability of detection  $\beta_{\min} = 0.95$ , then  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^+)} \approx 280/80 = 3.5$ , which means that the system  $\underline{\phi}^K$  designed to achieve  $\beta_{\min}$  under the least favorable emitter location distribution for  $\underline{\phi}^K$  has at most 3.5 times as many sensors as the system  $\underline{\phi}^K$  designed to achieve  $\beta_{\min}$  under any other distribution.



**Figure 8.1.** Probability of detection of the distributed detection system defined by (8.5) and (8.6) with the distribution of  $Y_0$  and  $k$  defined such that the probability of false alarm equals  $\alpha_{\max} = 0.05$  when  $L_e$  is distributed according to  $P_{L_e}^-$  that places the emitter in the boundary of a circular region of interest with probability 1,  $P_{L_e}^+$  that satisfies  $P[L_e = (0,0)] = 1$ , and the uniform distribution  $P_{L_e}^u$  for various values for the number of sensors  $K$ . In all the evaluations,  $\lambda_W = 1$ ,  $t_s = \lambda_W$ ,  $r_M = 15$ , and  $A_{\max} = 40$ .

Figure 8.1 also shows the probability of detection when  $L_e$  is uniformly distributed in  $S_e$ , illustrating that the probability of detection curve is always between the curves corresponding to the least and the most favorable distributions. For this example, Figure 8.1 shows that  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^u)} \approx 280/160 = 1.75$ .

It is important to emphasize that the above mentioned ratios can change significantly with the shape of the region of interest. If  $S_e$  is instead a square circumscribed by  $\partial B_{r_M}(0)$  and all the other parameters remain constant, then it is possible to obtain through numerical integration that  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^+)} \approx 6$  and  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^u)} \approx 3$ .

## 8.2 Large Sample Methods to Evaluate How Conservative the Least Favorable Distribution is

Although it is possible to obtain the ratios  $K^{(\Phi, P_{L_e}^-)} / K^{(\Phi, P_{L_e}^+)}$  and  $K^{(\Phi_{P_{L_e}^-|_{\mathcal{C}}}, P_{L_e}^-)} / K^{(\Phi_{P_{L_e}^+|_{\mathcal{C}}}, P_{L_e}^+)}$  numerically, it is also possible to obtain an asymptotic value for these ratios through the use of large sample methods.

The large sample method that will be used for this analysis is based on the *Asymptotic Relative Efficiency* (ARE). In this analysis, instead of considering the test of a simple hypothesis  $H_0 : \theta = 0$  against another simple hypothesis  $H_1 : \theta = \theta_1$ , one considers a sequence of detection systems  $\Phi = \{\underline{\phi}^n\}_{n=1}^\infty = \{(\{\phi_{i,n}, P_{Y_{i,n}}\}_{i=0}^n)\}_{n=1}^\infty$ , each one with  $n$  sensors for testing the simple hypothesis  $H_{0,n} : \theta = 0$  against another simple hypothesis  $H_{1,n} : \theta = \theta_{1,n} := c/\sqrt{n}$ , for any positive constant  $c$ . In general, such analysis can be accomplished when measurements have conditional density

functions with respect to a common measure that belong to a family of densities that is quadratic mean differentiable at  $\theta = 0$  [73].<sup>1</sup>

In the ARE analysis, one evaluates a sequence of detection systems  $\Phi$  through its *efficacy*. The efficacy of a sequence of detection systems determines the asymptotic probability of detection for a maximum asymptotic probability of false alarm as  $\theta_{1,n} \rightarrow 0$ . This notion is made precise by the following theorem, extracted from [73] in a simplified form:

**Theorem 8.1.** (Based on [73, Theorem 13.2.1])

Assume  $\{U_i\}_{i=1}^n$  are conditionally i.i.d. with conditional densities  $f_{U_i|\theta=0}(u)$  and  $f_{U_i|\theta=\theta_{1,n}}(u)$  with respect to a common measure,<sup>2</sup> and assume that such densities belong to a family of densities that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ . If there exists a function  $g : \mathbb{R} \rightarrow \mathbb{R}$  and a number  $\sigma > 0$  such that

1. The detection system  $\{\phi_{i,n}, P_{Y_{i,n}}\}_{i=0}^n$  rejects  $H_{0,n}$  when  $\sqrt{n}[\mathbf{T}_n(\mathbf{u}) - g(0)] > \tau_n$ , where  $\mathbf{T}_n(\mathbf{u})$  is some given statistic,  $\tau_n \rightarrow Q^{-1}(\alpha_{\max}) \cdot \sigma$  as  $n \rightarrow \infty$ ,  $\alpha_{\max} > 0$  is a desired asymptotic probability of false alarm, and  $Q(\cdot)$  is the complementary c.d.f. of a Gaussian random variable with zero mean and unit variance;
2.  $\sqrt{n}[\mathbf{T}_n(\mathbf{U}) - g(\theta)]$  converges in distribution to a Gaussian distribution with zero mean and variance  $\sigma^2$  as  $n \rightarrow \infty$  under either  $\theta = 0$  or  $\theta = c/\sqrt{n}$ , where  $c$  is a given positive constant;
3. The function  $g$  has positive right-hand derivative  $g'$  at 0;

then the probability of false alarm converges to  $\alpha_{\max}$  and the probability of detection converges to

$$Q\left(Q^{-1}(\alpha_{\max}) - c \cdot \frac{g'(0)}{\sigma}\right), \quad (8.8)$$

and the efficacy  $\zeta(\Phi)$  of the sequence of detectors  $\Phi$  is given by  $g'(0)/\sigma$ .

From (8.8), if two sequences of detectors have two different efficacies, then they have different asymptotic probabilities of detection.

In order to compare two sequences of detectors at the same asymptotic detection performance, one considers appropriate *subsequences* of such sequences to find the ARE metric. Instead of considering the two whole sequences

<sup>1</sup>A family of densities  $\{f_{Z|\theta=\theta^*}(\cdot) : \theta^* \in (-\varepsilon, \varepsilon)\}$  with respect to a common measure  $\mu$  is quadratic mean differentiable at  $\theta^* = 0$  if the Fréchet derivative of the map  $\theta^* \rightarrow \sqrt{f_{Z|\theta=\theta^*}(\cdot)}$  from  $\mathbb{R}$  to the function space  $L_2(\mu)$  exists at  $\theta^* = 0$  [73]; i.e., if there exists a function  $\eta(\cdot, 0)$  such that

$$\lim_{|h| \rightarrow 0} \frac{\left\| \sqrt{f_{Z|\theta=h}(\cdot)} - \sqrt{f_{Z|\theta=0}(\cdot)} - \eta(\cdot, 0) \cdot h \right\|_2}{|h|} = 0, \quad (8.7)$$

where  $\|\cdot\|_2$  represents the norm in the function space  $L_2(\mu)$ .

<sup>2</sup>This theorem applies to both centralized and distributed detection systems. In the case of centralized detection systems, let  $U_i = Z_i$ .

$\{\underline{\phi}_{(1)}^n\}_{n=1}^\infty$  and  $\{\underline{\phi}_{(2)}^n\}_{n=1}^\infty$  in which the number of sensors  $K = n$  for each  $n$ ; one considers two subsequences of such sequences:

$$\begin{aligned}\Phi_{(1)} &:= \{\underline{\phi}_{(1)}^{K^{(1)}(n)}\}_{n=1}^\infty, \\ \Phi_{(2)} &:= \{\underline{\phi}_{(2)}^{K^{(2)}(n)}\}_{n=1}^\infty,\end{aligned}$$

where  $K^{(1)}(n) := d^{(1)} \cdot n$  and  $K^{(2)}(n) := d^{(2)} \cdot n$  for appropriate values of  $d^{(1)}$  and  $d^{(2)}$  such that both subsequences have the same asymptotic detection performance, as explained in Appendix I.2.

The Asymptotic Relative Efficiency (ARE) metric of the sequence of detectors  $\Phi_{(2)}$  with respect to the sequence of detectors  $\Phi_{(1)}$  [73, p. 536] is then given by

$$\lim_{n \rightarrow \infty} \frac{K^{(1)}(n)}{K^{(2)}(n)} = \left( \frac{\zeta(\Phi_{(2)})}{\zeta(\Phi_{(1)})} \right)^2 \quad (8.9)$$

and provides the asymptotic ratio of the number of sensors required by each design to achieve a prescribed detection performance. Details of the derivation of this known result can be found in Appendix I.2.

The ARE metric has the following *product property*: if the system designer knows the ARE metric of  $\Phi_{(2)}$  with respect to  $\Phi_{(1)}$  and the ARE metric of  $\Phi_{(3)}$  with respect to  $\Phi_{(2)}$ , then the ARE metric of  $\Phi_{(3)}$  with respect to  $\Phi_{(1)}$  can be found by the product of the two ARE metrics; i.e.,

$$\lim_{n \rightarrow \infty} \frac{K^{(1)}(n)}{K^{(3)}(n)} = \left( \frac{\zeta(\Phi_{(2)}) \zeta(\Phi_{(3)})}{\zeta(\Phi_{(1)}) \zeta(\Phi_{(2)})} \right)^2. \quad (8.10)$$

Using the product property of the ARE metric, the ARE metric for several types of detection systems will be derived in the next subsections. For these derivations, the additive model (2.1) described in Chapter 2 and the following two cases of interest will be considered:

1. **Gaussian case:** For any  $i$ ,  $A_i = \theta \xi(\|L_i - L_e\|)$ , and  $W_i$  is Gaussian distributed with zero mean and known variance  $\text{var}[W] > 0$ .
2. **Poisson case:** For any  $i$ ,  $A_i$  conditioned on  $\{\theta = \theta_{1,n}, L_i = l_i, L_e = l_e\}$  is Poisson distributed with parameter  $\theta_{1,n} \xi(\|l_i - l_e\|)$  and  $W_i$  is Poisson distributed with a known parameter  $\lambda_W > 0$ .

Under the assumptions of Propositions 6.2 or 6.3, Propositions I.4 and I.5 in the appendix indicate that, when considering a centralized detection system, the resulting conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  from the Gaussian case or the resulting conditional p.m.f.s  $P_{Z_i|\theta=0}(z)$  and  $P_{Z_i|\theta=\theta_{1,n}}(z)$  from the Poisson case, both under  $L_e \in S_e^-$  or under  $L_e = (0, 0)$ , belong to families that are quadratic mean differentiable at  $\theta = 0$ . When considering a distributed detection system that uses the practical sensor function  $\phi_i(z_i, y) = \mathbf{1}\{z_i > t_s\}$ , Propositions I.6 and I.7 in the appendix indicate that the resulting conditional p.m.f.s  $P_{U_i|\theta=0}(z)$  and  $P_{U_i|\theta=\theta_{1,n}}(z)$  from both the Gaussian or the Poisson cases under  $L_e \in S_e^-$  or under  $L_e = (0, 0)$  belong to a family that is quadratic mean differentiable at  $\theta = 0$ .



### 8.2.1 The ARE of the Maximax Centralized Detection System with Respect to the Maximin Centralized Detection System

**Proposition 8.2.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption VII.*

*Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , define  $S_e^- := S_e \cap \partial B_{r_M}(0)$  and  $S_e^+ = \{0\}$ . Let the sensor locations  $\{L_i\}_{i=1}^K$  be i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  for some  $c > 0$ ;*

*Let  $P_{L_e}^-$  be a distribution that satisfies  $P[L_e \in S_e^-] = 1$  and assume a first sequence of centralized detection systems  $\Phi_{P_{L_e}^-} := \{\phi_{P_{L_e}^-}^n\}_{n=1}^\infty$  in which, for each  $n$ ,  $\phi_{P_{L_e}^-}^n$  is the centralized detection system with  $n$  sensors defined in Proposition 7.4 with  $L_e$  distributed according to  $P_{L_e}^-$ .*

*Let  $P_{L_e}^+$  be a distribution that satisfies  $P[L_e \in S_e^+] = 1$  and assume a second sequence of centralized detection systems  $\Phi_{P_{L_e}^+} := \{\phi_{P_{L_e}^+}^n\}_{n=1}^\infty$  in which, for each  $n$ ,  $\phi_{P_{L_e}^+}^n$  is the centralized detection system with  $n$  sensors defined in Proposition 7.4 with  $L_e$  distributed according to  $P_{L_e}^+$ .*

*Under either the Poisson or the Gaussian case,*

- *the sequence  $\Phi_{P_{L_e}^-}$  is a sequence of maximin detection systems and reaches the maximum efficacy when the distribution of  $L_e$  satisfies  $P[L_e \in S_e^-] = 1$ ;*
- *the sequence  $\Phi_{P_{L_e}^+}$  is a sequence of maximax detection systems and reaches the maximum efficacy when the distribution of  $L_e$  satisfies  $P[L_e \in S_e^+] = 1$ ; and*
- *the ARE of  $\Phi_{P_{L_e}^+}$  with respect to  $\Phi_{P_{L_e}^-}$  is given by*

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi_{P_{L_e}^-}, P_{L_e}^-)}(n)}{K^{(\Phi_{P_{L_e}^+}, P_{L_e}^+)}(n)} = \left( \frac{E[\xi(\|L_i - (0, 0)\|)]}{E[\xi(\|L_i - (0, r_M)\|)]} \right)^2. \quad (8.11)$$

*Proof:* The conclusions that  $\Phi_{P_{L_e}^-}$  is a sequence of maximin detection systems and  $\Phi_{P_{L_e}^+}$  is a sequence of maximax detection systems follows from Propositions 7.8 and 7.9. Observe that the same additive model, the same set of assumptions from Chapter 5, and the uniform distribution for sensor locations  $\{L_i\}_{i=1}^K$  required by these propositions are also assumed here. The conditions on  $\mathcal{Z}$  and  $f_W$  required by these propositions are satisfied for the Gaussian case because the measurements  $\{Z_i\}_{i=1}^K$  take values in  $\mathcal{Z} = \mathbb{R}$ ,  $f_W(w)$  is the Gaussian p.d.f. and  $f_W(w) > 0$  for all  $w \in \mathbb{R}$ , and  $f_W(w - a)/f_W(w)$  is nondecreasing in  $w \in \mathbb{R}$  for all  $a \geq 0$ . The conditions on  $\mathcal{Z}$  and  $f_W$  are also satisfied for the Poisson case because the measurements  $\{Z_i\}_{i=1}^K$  take values in  $\mathcal{Z} = \mathbb{Z}_+$ ,  $f_W(w)$  is the Poisson p.m.f., which satisfies  $f_W(w) > 0$  for all  $w \in \mathbb{Z}_+$ , and  $f_W(w - a)/f_W(w)$  is nondecreasing in  $w \in \mathbb{Z}_+$  for all  $a \in \mathbb{Z}_+$ . Lastly, the condition that  $\forall a_i, P[A_i > a | \theta = \theta_{1,n}, \xi(\|L_i - L_e\|) = x]$  be a nondecreasing function of  $x$  is satisfied in the Gaussian case because  $A_i = \theta_{1,n}x$ ; and is satisfied in the Poisson case as shown in Proposition F.6 in the appendix.

The conclusions that  $\Phi_{P_{L_e}^-}$  and  $\Phi_{P_{L_e}^+}$  reach the maximum efficacy follow directly from [73, Lemma 13.3.1], which shows that a sequence of Neyman–Pearson detectors in which the detector used for each  $n$  is the optimal fusion function to decide between  $H_{0,n}$  and  $H_{1,n}$  has the maximum possible efficacy, given by the square root of the Fisher Information associated with the family  $\{f_{Z_i|\theta=\theta^*}(z)|\theta^* \in (-\varepsilon, \varepsilon)\}$  at  $\theta = 0$ . It is now possible to apply [73, Lemma 13.3.1] because, for either the case in which the distribution of  $L_e$  satisfies  $P[L_e \in S_e^-] = 1$  or  $P[L_e \in S_e^+] = 1$ , the measurements  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$ , as shown in Propositions 7.8 and 7.9, and the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  under  $L_e \in S_e^-$  and under  $L_e = (0, 0)$  belong to a family that is quadratic mean differentiable at  $\theta = 0$  for either the Gaussian or Poisson cases, as shown in Propositions I.4 and I.5 in the appendix.

Since both  $\Phi_{P_{L_e}^-}$  and  $\Phi_{P_{L_e}^+}$  achieve the maximum possible efficacy under their respective distribution for  $L_e$ , the ARE of  $\Phi_{P_{L_e}^+}$  with respect to  $\Phi_{P_{L_e}^-}$  is given by  $I^+(0)/I^-(0)$ , where  $I^-(\theta)$  is the Fisher information associated with the family  $\{f_{Z_i|\theta=\theta^*}(z)|\theta^* \in (-\varepsilon, \varepsilon)\}$  under  $P[L_e \in S_e^-] = 1$  and  $I^+(\theta)$  is the Fisher information associated with the family  $\{f_{Z_i|\theta=\theta^*}(z)|\theta^* \in (-\varepsilon, \varepsilon)\}$  under  $P[L_e \in S_e^+] = 1$ .

For either the Gaussian or the Poisson cases, as shown in Propositions I.4 and I.5 in the appendix, it is possible to determine that  $I^-(0)$  is proportional to  $E[\xi(\|L_i - l_e\|)]^2$  for some  $l_e \in S_e^-$  and  $I^+(0)$  is proportional to  $E[\xi(\|L_i - l_e\|)]^2$  for  $l_e = 0$ , and (8.11) follows. ■

Proposition 8.2 gives conditions under which the asymptotic value for the looser bound of (8.4) can be simply obtained from the expected value of the amplitude function  $\xi$ .

In the next subsection, it is shown that the asymptotic value for the looser bound of (8.3) can be obtained from the expected value of the amplitude function  $\xi$  as well for an important practical centralized detection system.

## 8.2.2 The ARE of Practical Centralized Detection Systems

Although the ARE metric is usually used to compare two candidate designs, it is also possible to use the ARE metric to compare a given design under two different measurement distributions. Consider a given sequence of detectors  $\Phi$ . The efficacy of  $\Phi$  depends on the distribution of the measurements. Thus, if two different measurement distributions are being considered, it is possible to determine the ARE metric of  $\Phi$  under the first distribution with respect to  $\Phi$  under the second distribution.

This is precisely what is needed to compute the asymptotic value for the looser bound of (8.3); i.e., given a sequence of detectors  $\Phi$ , the goal is to compute the ARE metric of  $\Phi$  under a distribution  $P_{L_e}^+$  that is most favorable for each detection system in  $\Phi$ , with respect to the same  $\Phi$ , but now under a distribution  $P_{L_e}^-$  that is least favorable for each detection system in  $\Phi$ .

**Proposition 8.3.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V.*

*For each  $i$ , let  $Z_i$  assume values in  $\mathcal{Z}$ , with  $\mathcal{Z} = \mathbb{R}$ ,  $\mathcal{Z} = \mathbb{R}_+$ ,  $\mathcal{Z} = \mathbb{Z}$ , or  $\mathcal{Z} = \mathbb{Z}_+$ . Let  $\mathcal{Z}_+ = \mathcal{Z} \cap \mathbb{R}_+$ .*

Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , define  $S_e^- := S_e \cap \partial B_{r_M}(0)$  and  $S_e^+ = \{0\}$ . Let  $P_{L_e^-}$  denote a distribution that satisfies  $P[L_e \in S_e^-] = 1$  and let  $P_{L_e^+}$  denote the distribution that satisfies  $P[L_e \in S_e^+] = 1$ .

Consider the sequence of centralized detection systems  $\Phi^{(\Sigma)}$  in which, for each  $n$ , the centralized detection system with  $n$  sensors uses the fusion function

$$\phi_{0,n}^{\Sigma}(\mathbf{z}) := \mathbf{1} \left\{ \sum_{i=1}^n z_i > t_n \right\} \quad (8.12)$$

with  $t_n := \sqrt{n} \cdot \sqrt{\text{var } W} \cdot Q^{-1}(\alpha_{\max}) + n \cdot E[W]$  for a given  $\alpha_{\max} \in (0, 1)$ .

If the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  under either  $P_{L_e^-}$  or  $P_{L_e^+}$  belong to a family that is quadratic mean differentiable at  $\theta = 0$ , and if

- the sensor locations  $\{L_i\}_{i=1}^K$  are i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  for some  $c > 0$ ;
- for any  $a_i$ ,  $P[A_i > a_i | \theta = \theta_1, \xi(\text{dist}(L_i, L_e)) = x]$  is a nondecreasing function of  $x$ ;
- $E[A_i | \theta = \theta^*, L_e = l_e] = \theta^* \cdot E[\xi(\|L_i - l_e\|)]$ ;
- $E[W] < \infty$ ,  $\text{var } W < \infty$ ,  $E[A_i | \theta = \theta^*, L_e = l_e] < \infty$ , and  $\text{var}[A_i | \theta = \theta^*, L_e = l_e] < \infty$ ;
- $\text{var}[A_i | \theta = \theta^*, L_e = l_e] \rightarrow 0$  as  $\theta^* \rightarrow 0$ ; and
- $E[\xi(\|L_i - l_e\|)] > 0$ ;

for any  $\theta^* \in (0, 1)$  and any  $l_e \in \{S_e^- \cup (0, 0)\}$ , then the efficacies of  $\{\phi_{0,n}^{\Sigma}\}_{n=1}^{\infty}$  under  $P[L_e \in S_e^-] = 1$  and  $P[L_e \in S_e^+] = 1$  are respectively given by

$$E[\xi(\|L_i - (0, r_M)\|)] / \sqrt{\text{var } W}, \quad (8.13)$$

$$E[\xi(\|L_i - (0, 0)\|)] / \sqrt{\text{var } W}, \quad (8.14)$$

and the ARE of  $\Phi^{(\Sigma)}$  under  $P_{L_e^+}$  with respect to  $\Phi^{(\Sigma)}$  under  $P_{L_e^-}$  is given by

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi^{(\Sigma)}, P_{L_e^-})}(n)}{K^{(\Phi^{(\Sigma)}, P_{L_e^+})}(n)} = \left( \frac{E[\xi(\|L_i - (0, 0)\|)]}{E[\xi(\|L_i - (0, r_M)\|)]} \right)^2. \quad (8.15)$$

*Proof:* This proposition follows from the application of Theorem 8.1 to find the efficacy of the  $\Phi^{(\Sigma)}$  under  $P[L_e \in S_e^-] = 1$  and under  $P[L_e \in S_e^+] = 1$ . When applying this theorem, note that the random variables  $\{U_i\}_{i=1}^K$  of Theorem 8.1 are representing  $\{Z_i\}_{i=1}^K$  since centralized detection systems are being considered.

The condition that  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta$  follows from Proposition 7.7, which also gives that  $P_{L_e^-}$  and  $P_{L_e^+}$  are respectively a least and a most favorable distributions for each of the detection systems in the sequence of detection systems  $\Phi^{(\Sigma)}$ . Note that the conditions of Proposition 7.7 are satisfied because each detector in  $\Phi^{(\Sigma)}$  is in

$\mathcal{D}_c$ , as shown in Proposition 7.3, and all the other conditions of Proposition 7.7 are assumed here. The condition that the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  under either  $P_{L_e^-}$  or  $P_{L_e^+}$  belong to a family that is quadratic mean differentiable at  $\theta = 0$  is also assumed here.

Since  $\{Z_i\}_{i=1}^K$  does not depend on  $L_e$  when  $P[L_e \in S_e^-] = 1$ , let  $P_{L_e^-}$  be the distribution that satisfies  $P[L_e = (0, r_M)] = 1$  and note that it is possible to assume that  $(0, r_M) \in S_e^-$  when  $S_e$  is a regular convex polygon without loss of generality.

Define  $g(\theta) = E[Z_i|\theta, L_e = l_e]$  and observe that it satisfies condition 1 of Theorem 8.1. To see this, consider  $\mathbf{T}_n(\mathbf{z}) = (1/n) \sum_{i=1}^n z_i$ . Since  $E[A_i|\theta, L_e = l_e] = \theta E[\xi(\|L_i - l_e\|)]$ ,  $g(0) = E[W]$ . Thus, the sequence of detectors given by (8.12) is equivalent to the sequence of detectors specified in the theorem because  $\sum_{i=1}^n z_i > t_n \Leftrightarrow \sqrt{n}[(1/n) \sum_{i=1}^n z_i - E[W]] > \sqrt{n}[(t_n/n) - E[W]]$ .

Condition 2 follows from the Central Limit Theorem for the case  $\theta = 0$  and from the Lindeberg–Feller triangular array theorem [42] for the case  $\theta = c/\sqrt{n}$ . The specified assumptions on the function  $\xi(\cdot)$  and on the distributions of  $A_i$ ,  $W_i$ , and  $L_i$  are sufficient to satisfy the conditions of both theorems.

Condition 3 follows from the assumption  $E[A_i|\theta, L_e = l_e] = \theta \cdot E[\xi(\|L_i - l_e\|)]$ , which implies that  $g(\theta) = \theta \cdot E[\xi(\|L_i - l_e\|)] + E[W]$ , and from the assumption  $E[\xi(\|L_i - l_e\|)] > 0$ , which implies that  $g'(0) > 0$ .

Therefore, the efficacy of  $\Phi(\Sigma)$  under  $P[L_e \in S_e^-] = 1$  is given by  $E[\xi(\|L_i - (0, r_M)\|)]/\sqrt{\text{var} W}$  and the efficacy of  $\Phi(\Sigma)$  under  $P[L_e \in S_e^+] = 1$  is given by  $E[\xi(\|L_i - (0, 0)\|)]/\sqrt{\text{var} W}$ . By taking the square of the ratio of these efficacies, (8.15) is reached. ■

It is possible to show that the conditions on the distribution of  $A_i$  and  $W_i$  are satisfied for both the Gaussian and Poisson cases.

From Proposition 7.7, the distributions  $P_{L_e^-}$  and  $P_{L_e^+}$  defined are a least and a most favorable distribution for each detection system in  $\Phi(\Sigma)$ ; thus, the ARE metric obtained in (8.15) provides an upper bound for the number of additional sensors required by the centralized detection system given by (8.12) designed under  $P_{L_e^-}$  when compared to the number of sensors required by the centralized detection system given by (8.12) designed under the actual distribution for  $L_e$ .

Observing that the ARE metrics obtained in (8.11) and (8.15) are equal, it is possible to conclude that, when the conditions of the propositions are met, the same upper bound for the number of additional sensors required when considering  $P_{L_e^-}$  is achieved for both the case when the designer uses the optimum (maximin) centralized detection system and when the designer uses the practical centralized detection system that decides based on the sum of the measurements.

### 8.2.3 Bounds for the ARE Metric

The ARE metrics given by (8.11) and (8.15) can be determined numerically. Expressions for both  $E[\xi(\|L_i - (0, 0)\|)]$  and  $E[\xi(\|L_i - (0, r_M)\|)]$  in various cases of interest can be found in Appendix J. It is possible, however, to obtain simple lower and upper bounds for the ARE metric.

**Proposition 8.4.** *If  $L_i$  is uniformly distributed in a regular convex polygon with  $n$  vertices and circumscribed by  $B_{r_M}(0)$ , then for any nonnegative  $\xi$ ,*

$$4 \left( \frac{n}{n-2} \frac{\int_0^{r_M \cos(\pi/n)} \xi(r) r dr}{\int_0^{2r_M} \xi(r) r dr} \right)^2 \leq \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0,r_M)\|)]} \right)^2$$

$$\leq 4 \left( \frac{n}{n-2} \frac{\int_0^{r_M} \xi(r) r dr}{\int_0^{2r_M \sin^2(\pi/n)} \xi(r) r dr} \right)^2 \quad (8.16)$$

$$\leq 4 \left( \frac{n}{n-2} \right)^2 \text{ if } n \leq 4. \quad (8.17)$$

*Proof:* The bounds of this proposition follows from bounds on  $E[\xi(\|L_i - (0,0)\|)]$  and  $E[\xi(\|L_i - (0,r_M)\|)]$ .

For the bounds on  $E[\xi(\|L_i - (0,0)\|)]$ , use polar coordinates to write

$$E[\xi(\|L_i - (0,0)\|)] = \frac{1}{m(S_e)} \int_0^{2\pi} \int_0^{g_n(\omega)} \xi(r) \cdot r dr d\omega, \quad (8.18)$$

where  $g_n(\omega)$  defines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega) \in S_e$ . Since the regular convex polygon is circumscribed by  $B_{r_M}(0)$  and inscribed by  $B_{r_M \cos(\pi/n)}(0)$  [34],  $r_M \cos(\pi/n) \leq g_n(\omega) \leq r_M$ , and it follows that

$$\frac{2\pi}{m(S_e)} \int_0^{r_M \cos(\pi/n)} \xi(r) \cdot r dr \leq E[\xi(\|L_i - (0,0)\|)] \leq \frac{2\pi}{m(S_e)} \int_0^{r_M} \xi(r) \cdot r dr. \quad (8.19)$$

For the bounds on  $E[\xi(\|L_i - (0,r_M)\|)]$ ,

$$E[\xi(\|L_i - (0,r_M)\|)] = \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n(\omega)} \xi(r) \cdot r dr d\omega, \quad (8.20)$$

where  $h_n(\omega)$  defines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega)$  in the set  $S_e$  translated by  $(0, -r_M)$  and rotated by  $\pi$  as described in Appendix J. Since the translated and rotated  $S_e$  is circumscribed by  $B_{r_M}((0, r_M))$ ,  $h_n(\omega) \leq 2r_M$ ; and since the side of the regular convex polygon has length equal to  $2r_M \sin(\pi/n)$  [34] and the translated and rotated polygon contains the triangle formed by the two sides with the origin as common vertex and a line that connects the two non-common vertices of such sides,  $h_n(\omega)$  is always greater than the height of such triangle. Since the angle of the regular convex polygon is given by  $\pi - 2\pi/n$ ,  $h_n(\omega) \geq 2r_M \sin(\pi/n) \cos(\pi/2 - \pi/n) = 2r_M \sin^2(\pi/n)$ , and one can write

$$\frac{\pi - 2\pi/n}{m(S_e)} \int_0^{2r_M \sin^2(\pi/n)} \xi(r) \cdot r dr \leq E[\xi(\|L_i - (0,r_M)\|)]$$

$$\leq \frac{\pi - 2\pi/n}{m(S_e)} \int_0^{2r_M} \xi(r) \cdot r dr. \quad (8.21)$$

The upper and lower bounds of (8.16) follows from the bounds in (8.19) and (8.21).

When  $n \leq 4$ ,  $\sin^2(\pi/n) \geq 1/2$  and  $\int_0^{r_M} \xi(r) \cdot r dr \leq \int_0^{2r_M \sin^2(\pi/n)} \xi(r) \cdot r dr$  and (8.17) follows.  $\blacksquare$

It should be mentioned that the upper bound of (8.16) becomes loose as  $n$  increases; however, it may be considered satisfactory for  $n \leq 4$ .

**Proposition 8.5.** *If  $L_i$  is uniformly distributed in  $B_{r_M}(0) \subset \mathbb{R}^2$ , then for any nonnegative  $\xi$ ,*

$$\begin{aligned} 4 \left( \frac{\int_0^{r_M} \xi(r) r dr}{\int_0^{2r_M} \xi(r) r dr} \right)^2 &\leq \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0,r_M)\|)]} \right)^2 \\ &= \left( \frac{\int_0^{r_M} \xi(r) r dr}{\frac{1}{\pi} \int_0^{2r_M} \xi(r) r \cdot \left( \frac{\pi}{2} - \sin^{-1}\left(\frac{r}{2r_M}\right) \right) dr} \right)^2 \leq 9. \end{aligned} \quad (8.22)$$

*Proof:* For the equality in (8.22), use polar coordinates to write

$$E[\xi(\|L_i - (0,0)\|)] = \frac{2\pi}{\pi r_M^2} \int_0^{r_M} \xi(r) \cdot r dr, \quad (8.23)$$

$$E[\xi(\|L_i - (0,r_M)\|)] = \frac{1}{\pi r_M^2} \int_0^\pi \int_0^{2r_M \sin \omega} \xi(r) \cdot r dr d\omega, \quad (8.24)$$

as explained in Appendix J, and rewrite  $E[\xi(\|L_i - (0,r_M)\|)]$  as

$$\begin{aligned} E[\xi(\|L_i - (0,r_M)\|)] &= \frac{1}{\pi r_M^2} \int_0^\pi \int_0^{2r_M} \xi(r) \cdot r \cdot \mathbf{1}\{r \leq 2r_M \sin \omega\} dr d\omega \\ &= \frac{2}{\pi r_M^2} \int_0^{2r_M} \int_0^{\pi/2} \xi(r) \cdot r \cdot \mathbf{1}\{r \leq 2r_M \sin \omega\} d\omega dr \end{aligned} \quad (8.25)$$

$$\begin{aligned} &= \frac{2}{\pi r_M^2} \int_0^{2r_M} \int_0^{\pi/2} \xi(r) \cdot r \cdot \mathbf{1}\{\omega \geq \sin^{-1}(r/2r_M)\} d\omega dr \\ &= \frac{2}{\pi r_M^2} \int_0^{2r_M} \int_{\sin^{-1}(r/2r_M)}^{\pi/2} \xi(r) \cdot r d\omega dr \end{aligned} \quad (8.26)$$

where (8.25) follows from Fubini's theorem and the symmetry of the integrand, and the equality in (8.22) follows from (8.23) and (8.26).

For the lower bound of (8.22), recall that  $\xi(r) \geq 0$  and, from (8.24), write

$$E[\xi(\|L_i - (0,r_M)\|)] \leq \frac{1}{r_M^2} \int_0^{2r_M} \xi(r) \cdot r dr, \quad (8.27)$$

and the lower bound follows from (8.23) and from (8.24).

For the upper bound, a slightly different approach is taken. Let  $\mathcal{F} = \{(x,y) : \|(x,y)\| \leq r_M, y \leq 0, -\pi/3 < \arctan x/|y| \leq \pi/3\}$  and note that the closure of  $\mathcal{F}$  is the sector with central angle  $2\pi/3$ , symmetric in  $x$ , and with  $y \leq 0$ . It is now claimed that

$$E[\xi(\|L_i - (0,0)\|)] = 3 \cdot \int_{\mathcal{F}} \xi(\|l_i - (0,0)\|) dP_{L_i}(l_i). \quad (8.28)$$

To reach (8.28), let  $\mathcal{F}^{\circ 2\pi/3}$  be  $\mathcal{F}$  rotated by  $2\pi/3$  and  $\mathcal{F}^{\circ 4\pi/3}$  be  $\mathcal{F}$  rotated by  $4\pi/3$ . Note that  $\mathcal{F}$ ,  $\mathcal{F}^{\circ 2\pi/3}$ , and  $\mathcal{F}^{\circ 4\pi/3}$  are disjoint and  $B_{r_M}(0) = \mathcal{F} \cup \mathcal{F}^{\circ 2\pi/3} \cup \mathcal{F}^{\circ 4\pi/3}$ . Since the uniform distribution in  $B_{r_M}(0)$  is invariant to rotations, the integral of  $\xi(\|l_i\|)$  over  $\mathcal{F}$  is equal to the integral over any rotation of  $\mathcal{F}$ .

To reach the upper bound of (8.22), let  $\mathcal{F}^* = \mathcal{F} + (0, r_M)$  and note that

$$\begin{aligned} E[\xi(\|L_i - (0,r_M)\|)] &\geq \int_{\mathcal{F}^*} \xi(\|l_i - (0,r_M)\|) dP_{L_i}(l_i) \\ &= \int_{\mathcal{F}} \xi(\|l_i - (0,0)\|) dP_{L_i}(l_i), \end{aligned} \quad (8.29)$$

where (8.29) follows because the uniform distribution in  $B_{r_M}(0)$  is invariant to translations when the translated set remains within  $B_{r_M}(0)$ . ■

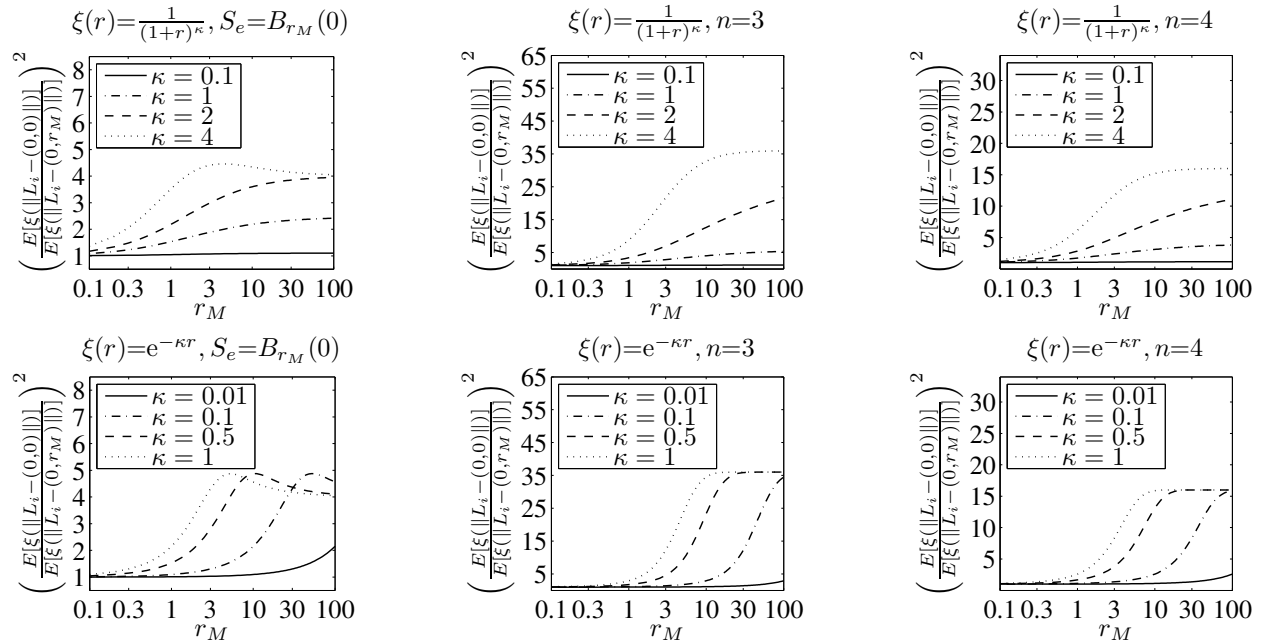
In both Propositions 8.4 and 8.5, if  $\xi$  is additionally nonincreasing, then it is possible to impose that the lower bound be always equal or greater than 1 because  $E[\xi(\|L_i - (0,0)\|)] \geq E[\xi(\|L_i - (0, r_M)\|)]$  in this case.

Proposition 8.5 gives an absolute upper bound that does not depend on  $\xi$  or  $r_M$ , and shows a system designer that, as the SNR decreases to zero, a sensor detection system designed considering  $\{L_i\}_{i=1}^K$  uniformly distributed in  $S_e = B_{r_M}(0)$  and considering the least favorable distribution for  $L_e$  would have at most 9 times the number of sensors that would be required if the actual distribution of  $L_e$  were known.

If the region of interest  $S_e$  is instead a regular convex polygon, the upper bound of Proposition 8.4 requires integration of  $\xi$ . The next section presents numerical evaluations of the ARE metric for both the circular and regular convex polygon regions.

## 8.2.4 Numerical Evaluations

Figure 8.2 presents the numerical evaluations of the resulting ARE of (8.11) for two types of amplitude functions with varying parameters for the cases in which the region of interest is a circular or a regular convex polygon. These figures were obtained through the numerical integration of formulas in Appendix J.



**Figure 8.2.** ARE of designs based on the most favorable distribution with respect to designs based on the least favorable distribution when  $\xi(r) = 1/(1+r)^\kappa$  and  $S_e = B_{r_M}(0)$  (top left),  $S_e$  is an equilateral triangle (top center), and  $S_e$  is a square (top right); and when  $\xi(r) = \exp(-\kappa \cdot r)$  and  $S_e = B_{r_M}(0)$  (bottom left),  $S_e$  is an equilateral triangle (bottom center), and  $S_e$  is a square (bottom right) for various values of  $\kappa$ .

The displays in this figure illustrate that, if the system is designed with  $r_M$  small enough such that the variation of the signal attenuation at any point of the region of interest is not significant, then the ARE of (8.11) is close to 1, showing that not much is lost when assuming a least favorable distribution. This is intuitive if one considers the extreme case in which the amplitude function does not decay with the distance; for example, when the value  $\kappa$  in the displays of the figure is close to 0. In the extreme case of  $\kappa = 0$ , the signal would be measured with the same amplitude at any sensor location, regardless of the emitter location and assuming the emitter located within the boundary of the region would not cause any impact.

However, if the variation of the signal attenuation is significant as the distance vary from 0 to  $r_M$ , then the ARE of (8.11) can be as high as 5 for the circular region of interest, as high as 36 for the triangular region of interest, or as high as 16 for the square region of interest.

Figure 8.2 suggests that the ARE metric converges to some value as  $r_M \rightarrow \infty$ . This observation is confirmed in the next subsection.

### 8.2.5 Asymptotic Bounds for the ARE Metric

**Proposition 8.6.** *If  $L_i$  is uniformly distributed in a regular convex polygon with  $n$  vertices and circumscribed by  $B_{r_M}(0)$ , then for any nonnegative  $\xi$  that satisfies  $\int_0^\infty \xi(r)r dr < \infty$ ,*

$$\lim_{r_M \rightarrow \infty} \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0,r_M)\|)]} \right)^2 = 4 \left( \frac{n}{n-2} \right)^2, \quad (8.30)$$

*and if  $L_i$  is uniformly distributed in  $B_{r_M}(0)$ , then*

$$\lim_{r_M \rightarrow \infty} \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0,r_M)\|)]} \right)^2 = 4. \quad (8.31)$$

*Proof:* The limit of (8.30) follows because both the lower and the upper bounds of Proposition 8.4 converge to the same limit. To see this, use the Monotone Convergence Theorem [47] to conclude that  $\int_0^{g(r_M)} \xi(r)r dr \rightarrow \int_0^\infty \xi(r)r dr$  for any increasing  $g(r_M)$  that converges to infinity as  $r_M \rightarrow \infty$ . Thus, the ratio of the integrals in both the lower and the upper bounds of Proposition 8.4 converge to 1.

To reach the limit of (8.31), recall (8.22) of Proposition 8.5 and write

$$\frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0,r_M)\|)]} = \frac{\int_0^{r_M} \xi(r)r dr}{\frac{1}{2} \int_0^{2r_M} \xi(r)r dr - \frac{1}{\pi} \int_0^{2r_M} \xi(r)r \sin^{-1}\left(\frac{r}{2r_M}\right) dr}. \quad (8.32)$$

From the Dominated Convergence Theorem [47], reach that  $\int_0^{2r_M} \xi(r)r \sin^{-1}\left(\frac{r}{2r_M}\right) dr \rightarrow 0$  and (8.32) converges to 2. ■

Although the value of  $r_M$  that defines the region of interest is always finite, the asymptotic bounds of Proposition 8.6 can be used as an approximation if the amplitude function is fast decaying, such that a sensor located at a distance on the order of  $r_M$  from the emitter senses the emitter with a signal level that can be approximated by 0.



## 8.2.6 The ARE Metric of a Maximin Centralized Detection System and a Practical Centralized Detection System Under a Least Favorable Distribution

While Section 8.2.1 presents the ARE of the maximax centralized detection system under a most favorable distribution with respect to the maximin centralized detection system under a least favorable distribution, and Section 8.2.2 presents the ARE of a practical centralized detection system under a most favorable distribution with respect to the same practical centralized detection system under a least favorable distribution, this subsection presents the ARE metric of the maximin centralized detection system with respect to a practical centralized detection system, both under a common distribution for the emitter location  $L_e$  that is least favorable for both the maximin and the practical centralized detection systems.

The results of this subsection can be used in two ways:

- From the product property of the ARE metric, the results of this subsection and subsection 8.2.1 allow a system designer to determine the ARE metric of the maximax centralized detection system under a most favorable distribution with respect to a practical centralized detection system under a least favorable distribution; and,
- The results of this subsection can answer the following design question: How many more sensors will a system designer need to deploy when assuming the least favorable distribution for a practical design instead of assuming a maximin design?

The next proposition shows that, under either the Gaussian or the Poisson cases, the ARE of a maximin design with respect to the practical centralized detection system that decides based on the sum of the measurements, both under a common distribution  $P_{L_e}$  that is least favorable for both of them, is equal to 1, which shows that the number of additional sensors required by the adoption of a simple practical design instead of the maximin design decreases to 0 as  $\theta_{1,n} \rightarrow 0$  and the number of sensors increases to infinity.

**Proposition 8.7.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V.*

*Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , define  $S_e^- := S_e \cap \partial B_{r_M}(0)$  and  $S_e^+ = \{0\}$ . Let the sensor locations  $\{L_i\}_{i=1}^K$  be i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  for some  $c > 0$ ;*

*Let  $P_{L_e}^-$  be a distribution that satisfies  $P[L_e \in S_e^-] = 1$  and let  $\Phi_{P_{L_e}^-} := \{\phi_{P_{L_e}^-}^n\}_{n=1}^\infty$  denote the sequence of centralized detection systems in which, for each  $n$ ,  $\phi_{P_{L_e}^-}^n$  is the centralized detection system with  $n$  sensors defined in Proposition 7.4 with  $L_e$  distributed according to  $P_{L_e}^-$ .*

*Let  $\Phi(\Sigma)$  denote the sequence of practical detection systems in which, for each  $n$ , the centralized detection system uses the fusion function  $\phi_{0,n}^\Sigma(\mathbf{z})$  of (8.12); i.e., the fusion function decides based upon the sum of the measurements.*

Under the Gaussian or the Poisson cases, the ARE metric of  $\Phi_{P_{L_e}^-}$  with respect to  $\Phi^{(\Sigma)}$  under  $P_{L_e}^-$  is equal to 1; i.e.,

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi^{(\Sigma)}, P_{L_e}^-)}(n)}{K^{(\Phi_{P_{L_e}^-}, P_{L_e}^-)}(n)} = 1. \quad (8.33)$$

*Proof:* First, observe that the conditions for Propositions 7.8 and 7.7 are satisfied, and it follows that the distribution  $P_{L_e}^-$  is a least favorable distribution for each detection system in either  $\Phi_{P_{L_e}^-}$  or  $\Phi^{(\Sigma)}$ . To see that the conditions for Propositions 7.8 and 7.7 are satisfied, observe that the same additive model, the same set of assumptions from Chapter 5, and the same conditions on  $S_e$ , on  $S_s$ , and on the distribution of  $\{L_i\}_{i=1}^K$  are assumed here. The detection systems in  $\Phi_{P_{L_e}^-}$  and  $\Phi^{(\Sigma)}$  are in  $\mathcal{D}_c$ . The conditions on  $\mathcal{X}$ ,  $f_W$ , and  $P[A_i > a | \theta = \theta_{1,n}, \xi(\|L_i - L_e\|) = x]$  required by these propositions are satisfied for either the Gaussian or the Poisson cases, as explained in Proposition 8.2.

To establish the conclusion, observe that, since the Gaussian and the Poisson cases satisfy the conditions of Proposition 8.3, the efficacy of the sequence of detectors  $\Phi^{(\Sigma)}$  is given by (8.13), which is equal to the square root of the Fisher Information at  $\theta = 0$  for both the Gaussian and Poisson cases, as shown in Propositions I.4 and I.5. Since the conditions of Proposition 8.2 are also satisfied,  $\Phi_{P_{L_e}^-}$  is a sequence of maximin detection systems with an efficacy also given by the square root of the Fisher Information at  $\theta = 0$ . Thus, the ARE metric of  $\Phi_{P_{L_e}^-}$  with respect to  $\Phi^{(\Sigma)}$  under the least favorable distribution equals 1. ■

The significance of Proposition 8.7 is that the penalty for using a least favorable distribution does not increase if the system designer adopts the practical centralized detection system that decides based upon the sum of the measurements in either the Gaussian or Poisson cases. In other words, although the ARE of the maximax centralized detection system with respect to the maximin centralized detection system can be as high as 9 for circular deployment regions and as high as 36 for triangular deployment regions, Proposition 8.7 shows that the ARE of the maximin centralized detection system with respect to such a practical detection system is 1; i.e., no additional sensors are required if the maximin centralized detection system is replaced by such a practical detection system. More generally, this means that, in either the Gaussian or Poisson cases, the system designer may adopt this simple practical detection system and be assured that it is asymptotically optimal under a least favorable distribution.

The conclusion of Proposition 8.7 may seem trivial because the sum of the measurements is a sufficient statistic for decision when deciding between  $H_{0,n} : Z_i = W_i$  and  $H_{1,n} : Z_i = a\theta_{1,n} + W_i$  when  $W_i$  are either Poisson or Gaussian distributed and  $a$  is a deterministic known value [65]; however, the sum of the measurements is *not* a sufficient statistic for decision when  $Z_i$  is distributed according to a mixture of Poisson or Gaussian distributions, which is the case considered here.

The next section contains the computation of the ARE metric of the practical centralized detection system considered in this section with respect to a practical distributed detection system under a common distribution that is least favorable for both of them. Note that, although it would be interesting to derive the ARE metric of the maximin centralized detection system with respect to the maximin distributed detection system, such a derivation is difficult

because maximin distributed detection systems may require sensor functions with different thresholds for any finite number of sensors, meaning that the  $\{U_i\}_{i=1}^K$  received at the fusion center will not be conditionally i.i.d., which is a necessary condition to apply Theorem 8.1.

## 8.2.7 The ARE Metric of a Practical Centralized Detection System and a Practical Distributed Detection System Under a Least Favorable Distribution

This section contains the computation of the ARE metric of the practical centralized detection system that decides based on the sum of measurements with respect to the practical distributed detection system that decides based on the sum of the binary outputs of sensor functions, both under a common distribution that is least favorable for both of them. Once this ARE metric is obtained, it is possible to use the product property of the ARE metric in order to obtain the ARE metric of the maximax or the maximin centralized detection system with respect to such a practical distributed detection system.

**Proposition 8.8.** *Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, VII, and VIII of Chapter 5. Let  $\text{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V.*

*Let the region of interest  $S_e$  be the closed ball  $B_{r_M}(0)$  or a regular convex polygon with circumscribing circle  $\partial B_{r_M}(0)$ , define  $S_e^- := S_e \cap \partial B_{r_M}(0)$  and  $S_e^+ = \{0\}$ . Let the sensor locations  $\{L_i\}_{i=1}^K$  be i.i.d. random variables with the uniform distribution in  $S_s = c \cdot S_e$  for some  $c > 0$ ;*

*Let  $\Phi(\Sigma^{dds})$  denote the sequence of practical distributed detection systems in which, for each  $n$ , the distributed detection system with  $n$  sensors uses the following fusion and sensor functions:*

$$\phi_{0,n}^{\Sigma,dds}(\mathbf{u}) = \mathbf{1} \left\{ \sum_{i=1}^n u_i > t_{0,n} \right\}, \quad (8.34)$$

$$u_i := \phi_{i,n}^{\Sigma,dds}(z_i) = \mathbf{1}\{z_i > t_s\}, \quad (8.35)$$

*for a given  $t_s$  such that  $U_i$  is not degenerate for any  $\theta \geq 0$ , and*

$$t_{0,n} := \sqrt{n} \cdot \sqrt{P[W > t_s](1 - P[W > t_s])} \cdot Q^{-1}(\alpha_{\max}) + n \cdot P[W > t_s], \quad (8.36)$$

*for some given  $\alpha_{\max} \in (0, 1)$ .*

*If*

- $Z_i$  is such that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  under  $L_e \in S_e^-$  and under  $L_e = (0, 0)$  belong to a family that is quadratic mean differentiable at  $\theta = 0$ ;
- $A_i \rightarrow 0$  in distribution as  $n \rightarrow \infty$ ;
- $P[A_i + W_i > t_s | \theta = \theta_{1,n}, L_i = l_i, L_e = l_e]$  is a differentiable function of the parameter  $x := \theta_{1,n} \cdot \xi(\|l_i - l_e\|)$  for any  $l_i \in S_s$ ,  $l_e \in S_e^- \cup S_e^+$ , and  $\theta_{1,n} \in [0, \varepsilon]$  for some  $\varepsilon > 0$ . Let  $h(t_s, x)$  be a such function; and

- the derivative of  $h(t_s, x)$  with respect to  $x$  at  $x = 0$  is positive; and there is a constant  $C$  that satisfies

$$\forall x \geq 0, \left| \frac{\partial}{\partial x} h(t_s, x) \right| \leq C; \quad (8.37)$$

then the efficacy of  $\Phi(\Sigma, dds)$  under  $P[L_e \in S_e^-] = 1$  is given by

$$\frac{\frac{\partial}{\partial x} h(t_s, x) \Big|_{x=0} \cdot E[\xi(\|L_i - (0, r_M)\|)]}{\sqrt{P[W > t_s](1 - P[W > t_s])}}; \quad (8.38)$$

and the efficacy of  $\Phi(\Sigma, dds)$  under  $P[L_e \in S_e^+] = 1$  is given by

$$\frac{\frac{\partial}{\partial x} h(t_s, x) \Big|_{x=0} \cdot E[\xi(\|L_i - (0, 0)\|)]}{\sqrt{P[W > t_s](1 - P[W > t_s])}}. \quad (8.39)$$

*Proof:* This proposition follows from the application of Theorem 8.1 to find the efficacy of  $\Phi(\Sigma, dds)$  under  $P[L_e \in S_e^-] = 1$  and under  $P[L_e \in S_e^+] = 1$ . The case  $P[L_e \in S_e^-] = 1$  is proven below and the case  $P[L_e \in S_e^+] = 1$  follows from the same reasoning.

Recall that  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d. under  $P[L_e \in S_e^-] = 1$ ; therefore, so are  $\{U_i\}_{i=1}^K$ . Furthermore, since  $\{Z_i\}_{i=1}^K$  and  $\{U_i\}_{i=1}^K$  do not depend on  $L_e$  when  $P[L_e \in S_e^-] = 1$ , pick  $l_e = (0, r_M)$  and, when  $S_e$  is a regular convex polygon, assume without loss of generality that such  $l_e$  is a vertex of the polygon.

Define  $g(\theta^*) = E[U_i | \theta = \theta^*, L_e = l_e]$  and  $\sigma = \sqrt{P[W_i > t_s](1 - P[W_i > t_s])}$ . It will be shown that such function and number satisfy the conditions of Theorem 8.1.

To see why condition 1 of Theorem 8.1 is satisfied, let  $\mathbf{T}_n(\mathbf{u}) = (1/n) \sum_{i=1}^n u_i$ , note that

$$\sum_{i=1}^n u_i > t_{0,n} \Leftrightarrow \sqrt{n} \left[ \frac{1}{n} \sum_{i=1}^n u_i - P[U_i = 1 | \theta = 0, L_e = l_e] \right] > \sqrt{n} \left[ \frac{t_{0,n}}{n} - P[U_i = 1 | \theta = 0, L_e = l_e] \right] =: \tau_n; \quad (8.40)$$

and, using  $t_{0,n}$  specified in (8.36), reach that the fusion function is on the form specified by Theorem 8.1 and  $\tau_n \rightarrow Q^{-1}(\alpha)\sigma$ .

Condition 2 follows from the Central Limit Theorem for the case  $\theta_{1,n} = 0$  and from the Lindeberg–Feller triangular array theorem [42] for the case  $\theta_{1,n} = c/\sqrt{n}$ . The conditions of both the Central Limit Theorem and the Lindeberg–Feller triangular array theorem are met since  $U_i \in \{0, 1\}$  and  $U_i$  is not degenerate.

To verify condition 3, observe that, for any  $\theta^* > 0$ ,  $\partial g(\theta^*)/\partial \theta^* = \partial P[A_i + W_i > t_s | \theta = \theta^*, L_e = l_e]/\partial \theta^*$ , and write

$$\begin{aligned} & \frac{\partial}{\partial \theta^*} P[A_i + W_i > t_s | \theta = \theta^*, L_e = l_e] \\ &= \frac{\partial}{\partial \theta^*} \int P[A_i + W_i > t_s | \theta = \theta^*, L_e = l_e, L_i = l_i] dP_{L_i}(l_i) \\ &= \int \frac{\partial}{\partial \theta^*} P[A_i + W_i > t_s | \theta = \theta^*, L_e = l_e, L_i = l_i] dP_{L_i}(l_i) \\ &= \int \frac{\partial}{\partial x} h(t_s, x) \Big|_{x=\theta^* \xi(\|l_i - l_e\|)} \xi(\|l_i - l_e\|) dP_{L_i}(l_i). \end{aligned} \quad (8.41)$$

$$= \int \frac{\partial}{\partial x} h(t_s, x) \Big|_{x=\theta^* \xi(\|l_i - l_e\|)} \xi(\|l_i - l_e\|) dP_{L_i}(l_i). \quad (8.42)$$

where (8.41) follows from Theorem I.1 and the assumptions on the derivative of  $h(t_s, x)$ ; and (8.42) follows from the chain rule. When  $\theta = 0$ , it follows that  $x = 0$ ; and since  $\partial h(t, 0)/\partial x$  is constant for all  $l_i$ , it follows that

$$\frac{\partial}{\partial \theta^*} g(0) = \frac{\partial}{\partial x} h(t_s, x) \Big|_{x=0} \cdot E[\xi(\|L_i - l_e\|)], \quad (8.43)$$

and, since it is assumed that  $\partial h(t, 0)/\partial x > 0$ , condition 3 of Theorem 8.1 is satisfied.  $\blacksquare$

Using the efficacies derived in Propositions 8.3 and 8.8, one obtains that the ARE of the practical centralized detection system that decides based on the sum of measurements (and satisfies the conditions of Proposition 8.3) with respect to the practical distributed detection system that decides based on the sum of the binary outputs of sensor functions (and satisfies conditions of Proposition 8.8), both under  $P_{L_e}^-$  that satisfies  $P[L_e \in S_e^-] = 1$  and is a least favorable distribution for both systems, is given by

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi(\Sigma, dds), P_{L_e}^-)}(n)}{K^{(\Phi(\Sigma), P_{L_e}^-)}(n)} = \frac{P[W > t_s](1 - P[W > t_s])}{\left(\frac{\partial}{\partial x} h(t_s, x) \Big|_{x=0}\right)^2 \cdot \text{var } W}, \quad (8.44)$$

The conditions of Proposition 8.8 are satisfied in either the Gaussian or the Poisson cases defined in Section 8.2, as shown in the next proposition.

**Proposition 8.9.** *Consider the same assumptions and distributed detection system defined in Proposition 8.8.*

*Under the Gaussian case defined in Section 8.2, the ARE of  $\Phi(\Sigma)$  of Proposition 8.3 with respect to  $\Phi(\Sigma, dds)$  of Proposition 8.8, both under  $P[L_e \in S_e^-] = 1$ , is given by*

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi(\Sigma, dds), P_{L_e}^-)}(n)}{K^{(\Phi(\Sigma), P_{L_e}^-)}(n)} = \frac{P[W > t_s](1 - P[W > t_s])}{f_W(t_s)^2 \cdot \text{var } W}, \quad (8.45)$$

where  $f_W(t_s)$  is the Gaussian p.d.f. with zero mean and variance  $\text{var } W$ .

*Under the Poisson case defined in Section 8.2, the resulting ARE of  $\Phi(\Sigma)$  with respect to  $\Phi(\Sigma, dds)$ , both under  $P[L_e \in S_e^-] = 1$ , is also given by (8.45) with  $f_W(t_s)$  representing the Poisson p.m.f. of parameter  $\lambda_W$  applied at the highest integer equal or lower than  $t_s$ .*

*Proof:* The result follows once it is shown that both the Gaussian and Poisson cases satisfy the conditions of Proposition 8.8 and compute the value of  $\partial h(t_s, x)/\partial x$  at  $x = 0$ .

For both Gaussian and Poisson cases, it is shown in Propositions I.6 and I.7 that the resulting conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  under  $L_e \in S_e^-$  and under  $L_e = (0, 0)$  belong to a family that is quadratic mean differentiable at  $\theta = 0$ .

Considering the Gaussian case, given  $\{\theta = \theta_{1,n}, L_i = l_i, L_e = l_e\}$ ,  $A_i = x := \theta_{1,n} \xi(\|l_i - l_e\|)$ . It is possible to see that  $A_i \rightarrow 0$  pointwise as  $\theta_{1,n} \rightarrow 0$ ;  $P[A_i + W_i > t_s | \theta = \theta_{1,n}, L_i = l_i, L_e = l_e] = Q((t_s - x)/\sigma_W)$ , which is a differentiable function of  $x$ ; its derivative with respect to  $x$  equals  $f_W(t_s - x)$ , which is bounded for all  $x$ ; and its derivative at  $x = 0$  equals  $f_W(t_s)$ , which is always positive.

Considering the Poisson case, given  $\{\theta = \theta_{1,n}, L_i = l_i, L_e = l_e\}$ ,  $A_i$  has a Poisson distribution with parameter  $x := \theta_{1,n} \xi(\|l_i - l_e\|)$ , which converges in distribution to 0 as  $\theta_{1,n} \rightarrow 0$ . Also, given  $\{\theta = \theta_{1,n}, L_i = l_i, L_e = l_e\}$ ,  $A_i + W_i$  has a Poisson distribution with parameter  $\lambda_W + x$ . Thus,

$$P[A_i + W_i > t_s | \theta = \theta_{1,n}, L_i = l_i, L_e = l_e] = 1 - \sum_{j=0}^{\lfloor t_s \rfloor} \frac{(\lambda_W + x)^j \exp(-\lambda_W - x)}{j!} =: h(t_s, x);$$

its derivative with respect to  $x$  is given by

$$\frac{\partial}{\partial x} h(t_s, x) = \exp(-\lambda_W - x) + \sum_{j=1}^{\lfloor t_s \rfloor} \frac{(\lambda_W + x)^j \exp(-\lambda_W - x)}{j!} \cdot \left(1 - \frac{j}{(\lambda_W + x)}\right),$$

which is a differentiable function of  $x$  for any  $\theta_{1,n}, l_i, l_e$ , and

$$\left| \frac{\partial}{\partial x} h(t_s, x) \right| \leq \exp(-\lambda_W) + 1 + (\lfloor t_s \rfloor / \lambda_W);$$

and the derivative at  $x = 0$  equals the Poisson p.m.f. of parameter  $\lambda_W > 0$  applied at the highest integer equal or lower than  $t_s$ ; i.e.,  $f_W(\lfloor t_s \rfloor)$ , which is always positive. ■

The efficacy of the practical distributed detection system defined in Proposition 8.8 also allows one to derive the ARE of the practical distributed detection system under  $P_{L_e}^+$  with respect to itself under  $P_{L_e}^-$ . More precisely, under the conditions of Proposition 8.8,

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi(\Sigma, dds), P_{L_e}^-)}(n)}{K^{(\Phi(\Sigma, dds), P_{L_e}^+)}(n)} = \left( \frac{E[\xi(\|L_i - (0, 0)\|)]}{E[\xi(\|L_i - (0, r_M)\|)]} \right)^2, \quad (8.46)$$

which is equal to the ARE of the practical centralized detection system of Proposition 8.3 under a most favorable distribution with respect to itself under a least favorable distribution. It should be mentioned however that this observation is not enough to conclude that the number of additional sensors incurred by the adoption of a least favorable distribution is independent of the type of detection system adopted; i.e., for some other detection system, the ARE of such a system under a most favorable distribution with respect to itself under a least favorable distribution may not equal (8.46).

Observe in (8.44) that the ARE of the practical centralized detection system  $\Phi(\Sigma)$  with respect to the practical distributed detection system  $\Phi(\Sigma), dds$ , both under  $P_{L_e}^-$ , does not depend on  $E[\xi(\|L_i - l_e\|)]$  for either the Gaussian or Poisson cases since  $P_{L_e}^-$  affects both efficacies in the same manner.

Instead, the ARE of  $\Phi(\Sigma)$  with respect to  $\Phi(\Sigma), dds$ , both under  $P_{L_e}^-$ , depends on the threshold  $t_s$  used in the sensor functions. Therefore, a system designer who wishes to use the practical distributed detection system of (8.34) and (8.35) may be interested in determining the threshold  $t_s^*$  that minimizes the ARE of  $\Phi(\Sigma)$  with respect to  $\Phi(\Sigma), dds$ .

The numerical determination of the threshold  $t_s^*$  that minimizes the ARE of  $\Phi(\Sigma)$  with respect to  $\Phi(\Sigma), dds$  should not impose difficulties; however, it is possible to determine  $t_s^*$  analytically in some cases.

In Appendix I.8, it is shown that  $t_s^* = 0$  for the Gaussian case, which means that, for this case,

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi(\Sigma_{dds}), P_{Le}^-)}(n)}{K^{(\Phi(\Sigma), P_{Le}^-)}(n)} = \frac{\pi}{2} \approx 1.57 \quad (8.47)$$

when sensor functions use  $t_s = 0$ .

### 8.2.8 Relationship to Prior Research

The results of Propositions 8.8 and 8.9 extend previous results that show that the efficacy of the practical distributed detection system defined in Proposition 8.8 to decide between  $H_{0,n} : Z_i = W_i$  and  $H_{1,n} : Z_i = a\theta_{1,n} + W_i$ , for a known deterministic value  $a$  and  $\{W_i\}$  i.i.d. with a Gaussian distribution of zero mean and variance  $\text{var } W$ , is given by  $f_W(t_s)/\sqrt{P[W > t_s](1 - P[W > t_s])}$  [37, 129]. Propositions 8.8 and 8.9 extend this result to the case  $H_{1,n} : Z_i = A_i + W_i$  when  $A_i$  is a random variable that depends on the distribution of  $L_e$ .

Regarding (8.47), a similar result was obtained in a simpler setting by the authors in [129] and [37], which considered deciding between  $H_0 : Z_i = W_i$  and  $H_1 : Z_i = a\theta_{1,n} + W_i$ , where the  $\{W_i\}_{i=1}^K$  are Gaussian random variables and  $a$  is a known deterministic value. The factor of  $\pi/2$  can also be found when evaluating the practical distributed detection system in the Bayesian setting: In [128], the authors claimed that  $t_s^* = 0$  is the threshold that minimizes the probability of error when deciding between  $H_0 : Z_i = W_i - a$  and  $H_1 : Z_i = W_i + a$ , where  $W_i$  are Gaussian random variables and  $a$  is a known deterministic value. The authors of [128] mentioned that the asymptotic probability of error converges to  $Q(\sqrt{2a^2/(\pi \text{var } W)})$  for the practical distributed detection system and converges to  $Q(\sqrt{a^2/\text{var } W})$  for the practical centralized detection system as  $K \rightarrow \infty$ ; and they highlighted the presence of the  $\pi/2$  factor causing degradation in the probability of error of the practical distributed detection system. The authors in [37] provided sufficient conditions for having  $t_s^* = 0$  minimize  $P[W > t_s](1 - P[W > t_s])/f_W(t_s)^2$ ; however, the conditions are not easily verifiable and the authors did not prove that such conditions are satisfied by the Gaussian distribution. It should also be mentioned that having  $t_s = 0$  in the practical distributed detection system of Proposition 8.8 forms the *sign test* in nonparametric settings, which is UMP when testing  $H_0 : P[Z_i > 0] = 1/2$  against  $H_1 : P[Z_i > 0] > 1/2$  [92].

### 8.2.9 Summary of Analysis

The results of this section allow a system designer to determine an approximate bound on the additional number of sensors necessary to attain a prescribed detection performance as the signal-to-noise ratio goes to zero when adopting a particular design under a least favorable distribution for the system in either the Gaussian or the Poisson cases. The results should be considered approximate because large sample methods are asymptotic.

Using the product property of the ARE metric and assuming the conditions of Propositions 8.2, 8.7, and 8.9, which include having either the Gaussian or the Poisson cases, one obtains the following:

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi_{P_{Le}^-}, P_{Le}^-)}(n)}{K^{(\Phi_{P_{Le}^+}, P_{Le}^+)}(n)} = \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0, r_M)\|)]} \right)^2 \quad (8.48)$$

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi^{(\Sigma)}, P_{Le}^-)}(n)}{K^{(\Phi_{P_{Le}^+}, P_{Le}^+)}(n)} = \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0, r_M)\|)]} \right)^2 \quad (8.49)$$

$$\lim_{n \rightarrow \infty} \frac{K^{(\Phi^{(\Sigma, dds)}, P_{Le}^-)}(n)}{K^{(\Phi_{P_{Le}^+}, P_{Le}^+)}(n)} = \left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0, r_M)\|)]} \right)^2 \cdot \frac{P[W > t_s](1 - P[W > t_s])}{f_W(t_s)^2 \cdot \text{var } W}, \quad (8.50)$$

where the reader should recall that  $\Phi_{P_{Le}^-}$  represents the maximin centralized detection system,  $\Phi_{P_{Le}^+}$  represents the maximax centralized detection system,  $\Phi^{(\Sigma)}$  represents the practical centralized detection system that decides based on the sum of the measurements (as defined in Proposition 8.3),  $\Phi^{(\Sigma, dds)}$  represents the practical distributed detection system that decides based on the sum of binary sensor outputs that equal 1 if the measurement is above a threshold  $t_s$  and 0 otherwise (as defined in Proposition 8.8), and  $P_{Le}^-$  and  $P_{Le}^+$  respectively represent least and most favorable distributions for the detection systems.

The expressions (8.48), (8.49), and (8.50) show that, at least for the Gaussian and Poisson cases, the number of additional sensors required when adopting a least favorable distribution for either the practical centralized or distributed detection system can be separated in two factors: the first one depends on the amplitude function  $\xi(\cdot)$  and the dimensions and shape of the region of interest and the deployment region; and the second one depends on the noise distribution and the threshold  $t_s$  used in the sensor functions.

The first factor is imposed by the least favorable distribution assumption, as shown by Proposition 8.3 and by expression (8.46). It was shown that this factor can be as high as 9 for regions of interest that form a disk, as high as 16 for square regions of interest, and as high as 36 for regions of interest that form an equilateral triangle; and it was shown that, as the sizes of the region of interest grow to infinity, this term converges to 4 for regions of interest that form a disk, 16 for square regions of interest, and 36 for regions of interest that form an equilateral triangle for any amplitude function that satisfies the conditions indicated in Proposition 8.6.

The second factor is imposed by the type of system: centralized or distributed. If centralized, it was shown in Proposition 8.7 that an additional number of sensors is not required when replacing the maximin centralized detection system with the practical centralized detection system that decides based on the sum of the measurements. If distributed, it was shown that, when the noise distribution is Gaussian with zero mean, the additional number of sensors incurred by the distributed detection system can be minimized by choosing  $t_s = 0$ , and the second factor becomes equal to  $\pi/2$ .



## Chapter 9

### Larger Deployment Regions

In order to improve the performance of a sensor detection system that is designed according to its least favorable distribution, this section explores the idea of larger sensor deployment regions; more precisely, this section compares the performance of two sensor deployment options:

- **Option 1:** Sensors randomly deployed in a deployment region  $S_s^{(1)} = S_e$ , and
- **Option 2:** Sensors randomly deployed in an *enlarged deployment region*  $S_s^{(2)} = cS_e$  for some  $c > 1$ ;

in which, for both options, sensors are uniformly distributed in the deployment region and the system designer adopts a least favorable distribution for the system. The focus will be on systems and cases that satisfy the conditions of the various propositions in Chapters 6 and 7 such that a least favorable distribution for the system is such that the emitter location  $L_e$  is distributed on  $S_e \cap \partial B_{r_M}(0)$  and the measurements under such a distribution are conditionally i.i.d. given  $\theta$ . Since such propositions include the case  $c > 1$ , meaning that a least favorable distribution for a system when  $S_s = S_e$  is also a least favorable distribution for the system when  $S_s = cS_e$  with  $c > 1$ , both deployment options are evaluated under a common least favorable distribution.<sup>1</sup>

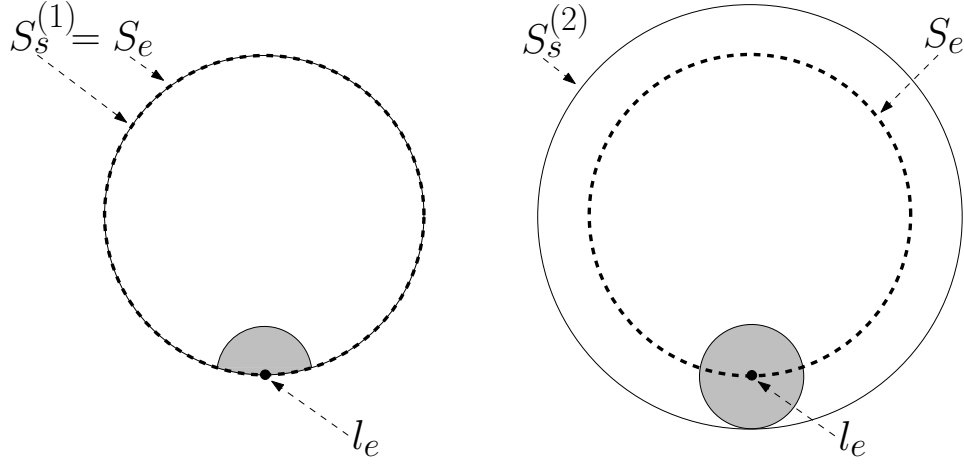
Deploying sensors in regions larger than the region of interest may seem counterintuitive since it proposes the deployment of sensors in areas that are known to not contain the emitter; however, the next example illustrates that such a strategy may indeed improve the detection performance when the emitter location is distributed according to a least favorable distribution for a system.

#### 9.1 Motivation for Larger Sensor Deployment Regions

Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let the signal  $A_i = \theta_1 \xi(\|L_i - L_e\|)$  with  $\xi(r) = \mathbf{1}\{r < r_0\}$  for some  $r_0 \in (0, 2r_M)$  and  $\theta_1 = 1$ . Let the noise distribution  $F_W$  be increasing.

---

<sup>1</sup>Throughout this chapter, the superscripts or subscripts (1) and (2) refer to the deployment options. Whenever the superscript or subscript is not used, the expression is intended for both options.



**Figure 9.1.** Two options for the sensor deployment region:  $S_s^{(1)} = S_e$  (left) and  $S_s^{(2)} = cS_e$  with  $c = 1 + r_0/r_M$  (right). The shaded region represents the region in which  $A_1 = 1$  when the emitter is located at  $(0, -r_M)$ .

Assume that  $S_e = B_{r_M}(0)$  and, as illustrated in Figure 9.1, consider two options for the sensor deployment region:  $S_s^{(1)} = S_e$ , and  $S_s^{(2)} = cS_e$  with  $c = 1 + r_0/r_M$ .

In either one of the options, consider a single sensor ( $K = 1$ ) and assume that its location  $L_1$  is uniformly distributed in  $S_s^{(j)}$  for  $j \in \{1, 2\}$ . Consider a centralized detection system in which the fusion function is given by  $\phi_0(z_1, y_0) = \mathbf{1}\{z_1 > t_0\}$ . Let  $t_0$  be such that  $P[Z_1 > t_0 | \theta = 0] \leq \alpha_{\max}$  for some  $\alpha_{\max} \in (0, 1)$ . Note that the probability of false alarm does not depend on the sensor deployment option.

Under these conditions and under either of the deployment options, if the emitter location  $L_e$  has a distribution  $P_{L_e}^{(l_e^-)}$  that satisfies  $P[L_e = l_e^-] = 1$  for  $l_e^- = (0, -r_M)$ , then the conditions of Proposition 7.7 are satisfied and  $P_{L_e}^{(l_e^-)}$  is a least favorable distribution for  $\phi_0$  for either deployment option.

As shown in Appendix G, the resulting probability of detection  $\beta(P_{L_e}^{(l_e^-)}, \phi_0)$  for either deployment option  $j \in \{1, 2\}$  satisfies

$$\begin{aligned} \beta(P_{L_e}^{(l_e^-)}, \phi_0) &= P[L_1 \in (S_s^{(j)} \cap B_{r_0}^\circ(l_e^-))] \cdot [F_W(t_0) - F_W(t_0 - 1)] + (1 - F_W(t_0)) \\ &= \frac{m(S_s^{(j)} \cap B_{r_0}^\circ(l_e^-))}{m(S_s^{(j)})} \cdot [F_W(t_0) - F_W(t_0 - 1)] + (1 - F_W(t_0)). \end{aligned} \quad (9.1)$$

Since  $F_W(t_0) - F_W(t_0 - 1) > 0$ , a larger sensor deployment region (option 2) provides a higher probability of detection than  $S_s = S_e$  (option 1) if

$$\frac{m(S_s^{(2)} \cap B_{r_0}^\circ(l_e^-))}{m(S_s^{(2)})} > \frac{m(S_s^{(1)} \cap B_{r_0}^\circ(l_e^-))}{m(S_s^{(1)})} \quad (9.2)$$

and vice-versa.

Since  $S_s^{(1)} \cap B_{r_0}^\circ(l_e^-)$  contains the sector  $[\sin^{-1}(r_0/(2r_M)), \pi - \sin^{-1}(r_0/(2r_M))]$  of  $B_{r_0}^\circ(l_e^-)$ ,<sup>2</sup> and  $(S_s^{(1)} \cap B_{r_0}^\circ(l_e^-)) \subset B_{r_0}^\circ(l_e^-) \cap \{(x, y) : y \geq -r_M\}$ ; i.e.,  $S_s^{(1)} \cap B_{r_0}^\circ(l_e^-)$  is a subset of the half-ball  $B_{r_0}^\circ(l_e^-)$ , one obtains

$$\frac{r_0^2}{\pi r_M^2} \cdot \left[ \frac{\pi}{2} - \sin^{-1}\left(\frac{r_0}{2r_M}\right) \right] < \frac{m(S_s^{(1)} \cap B_{r_0}^\circ(l_e^-))}{m(S_s^{(1)})} < \frac{r_0^2}{2r_M^2}. \quad (9.3)$$

Since  $S_s^{(2)} \cap B_{r_0}^\circ(l_e^-) = B_{r_0}^\circ(l_e^-)$ ,

$$\frac{m(S_s^{(2)} \cap B_{r_0}^\circ(l_e^-))}{m(S_s^{(2)})} = \frac{\pi r_0^2}{\pi (r_0 + r_M)^2} = \left( \frac{r_0}{r_0 + r_M} \right)^2. \quad (9.4)$$

Using (9.2) and the bounds of (9.3) and (9.4), if  $(r_0 + r_M)^2(\pi/2 - \sin^{-1}(r_0/(2r_M))) \geq \pi r_M^2$ , then deploying in  $S_s^{(1)}$  offers the higher probability of detection; however, if  $r_0 \leq (\sqrt{2} - 1)r_M$ , then deploying in the enlarged region  $S_s^{(2)}$  offers the higher probability of detection.

This example illustrates that, although a larger sensor deployment region reduces the sensor density within the region of interest  $S_e$ , which is reflected by the larger area of  $S_s^{(2)}$  in Figure 9.1, it increases the number of sensors that can potentially sense the emitter at a close range, which is reflected by the larger area of the shaded region in the right graph of Figure 9.1.

Although this example considers a single-sensor detection system in simple conditions, it illustrates that determining which deployment option provides better performance is not a trivial task; i.e., deploying in enlarged regions may or may not improve the detection performance of a system.

In the next section, the analysis is extended to more general systems and conditions while considering a low SNR regime and a minimum detection requirement. Since these considerations imply a large number of sensors, the large sample method based on the asymptotic relative efficiency (ARE) metric is adopted.

## 9.2 Comparing Deployment Options using the ARE Metric

Similarly as in Section 8.2, the ARE metric will be used to compare the performance of a given detection system under either of the deployment options. More precisely, consider a sequence of detection systems  $\Phi$  that satisfies the conditions of Theorem 8.1. The performance of each detection system in  $\Phi$  depends on the distribution of  $\{L_i\}_{i=1}^K$ . For  $j \in \{1, 2\}$ , let  $\Phi_{(j)}$  denote the sequence  $\Phi$  when the sensors are distributed according to the deployment option  $j$ . The two deployment options will be compared using the ARE metric of  $\Phi_{(2)}$  with respect to  $\Phi_{(1)}$ , which is given by

$$\lim_{n \rightarrow \infty} \frac{K^{(1)}(n)}{K^{(2)}(n)} = \left( \frac{\zeta(\Phi_{(2)})}{\zeta(\Phi_{(1)})} \right)^2, \quad (9.5)$$

<sup>2</sup>To see this, translate  $S_s^{(1)}$  and  $B_{r_0}^\circ(l_e^-)$  by  $(0, r_M)$  and note that the translated  $S_s^{(1)}$  is given in polar coordinates by  $\{(r, \omega) : r \leq 2r_M \sin \omega, \omega \in [0, \pi]\}$ , as shown in Appendix J. Pick any location  $l^*$  within the sector  $[\sin^{-1}(r_0/(2r_M)), \pi - \sin^{-1}(r_0/(2r_M))]$  of the translated  $B_{r_0}^\circ(l_e^-)$ ; i.e., its polar coordinates satisfy  $r^* < r_0$  and  $\omega^* \in [\sin^{-1}(r_0/(2r_M)), \pi - \sin^{-1}(r_0/(2r_M))]$ . This implies having  $\sin(\omega^*) \geq r_0/(2r_M)$ , which means that  $2r_M \sin(\omega^*) \geq r_0 > r^*$  and one concludes that  $l^*$  belongs to  $S_s^{(1)}$ .

where  $K^{(1)}(n)$  and  $K^{(2)}(n)$  represent the number of sensors required in each deployment option such that they have the same asymptotic detection performance, and  $\zeta(\Phi_{(1)})$  and  $\zeta(\Phi_{(2)})$  represent the efficacies of each deployment option.

In order to compute the ARE metric of  $\Phi_{(2)}$  with respect to  $\Phi_{(1)}$ , the results of Propositions 8.2, 8.3, and 8.8 of Chapter 8 will be used. For that, throughout this section, it is assumed that the conditions for either one of these propositions are satisfied, which means that either the Gaussian or Poisson cases defined in Chapter 8 is assumed; the sequence of sensor detection systems  $\Phi$  is considered to be either

- the maximin centralized detection system, as in Proposition 8.2;
- the practical centralized detection system that decides based on the sum of measurements, as in Proposition 8.3;
- or
- the practical distributed detection system that decides based on the sum of binary sensor outputs, as in Proposition 8.8;

and it is assumed that the emitter location is distributed according to  $P_{L_e}^-$  that satisfies  $P[L_e \in S_e^-] = 1$  for  $S_e^- := S_e \cap \partial B_{r_M}(0)$ . Recall that, from Propositions 8.2, 8.3, and 8.8,  $P_{L_e}^-$  is a least favorable distribution for any detection system in any of the sequences of detection systems  $\Phi$  identified above. Further observe that the conditions of such propositions include having  $L_i$  being uniformly distributed in  $S_s = cS_e$  for any  $c > 0$ , which means that the conclusions of such propositions are applicable to either the deployment options 1 or 2, which considers  $c > 1$ .

From the conclusions of Propositions 8.2, 8.3, and 8.8, the efficacies of either one of the identified sequences of detection systems  $\Phi$  under  $P_{L_e}^-$  depend on the distribution of  $L_i$  only through the factor  $E[\xi(\|L_i - (0, r_M)\|)]$ . Therefore, for either of these detection sequences, the ARE of the deployment option 2 with respect to the deployment option 1 is given by

$$ARE_{21} := \lim_{n \rightarrow \infty} \frac{K^{(1)}(n)}{K^{(2)}(n)} = \left( \frac{E[\xi(\|L_i^{(2)} - (0, r_M)\|)]}{E[\xi(\|L_i^{(1)} - (0, r_M)\|)]} \right)^2, \quad (9.6)$$

where  $L_i^{(j)}$  represents the location of sensors when deployed according to option  $j$ .

### 9.3 Sufficient Conditions to Adopt Deployment Option 2

Determining whether  $ARE_{21} > 1$  for a given amplitude function  $\xi(\cdot)$  would likely require the numerical evaluation of  $E[\xi(\|L_i^{(1)} - (0, r_M)\|)]$  and  $E[\xi(\|L_i^{(2)} - (0, r_M)\|)]$  for various  $c$ . Expressions for these quantities can be found in Appendix J. Although the numerical evaluation of such expressions should not impose significant challenges, it is possible to find sufficient conditions to have  $ARE_{21} > 1$  in some cases.<sup>3</sup>

<sup>3</sup>A note about the notation: from here to the end of this chapter,  $n$  denotes the number of vertices of a polygon, as opposed to the use in (9.5) and (9.6).

**Proposition 9.1.** Suppose that  $S_e = B_{r_M}(0)$  or  $S_e$  is a regular convex polygon with  $n$  vertices circumscribed by  $B_{r_M}(0)$ . Let  $L_i^{(1)}$  be uniformly distributed in  $S_s = S_e$  and let  $L_i^{(2)}$  be uniformly distributed in  $S_s = cS_e$  for some  $c \in (1, \sqrt{2})$ . If  $\xi(r) \geq 0$  is a decreasing function and

$$\lim_{r \rightarrow \infty} \frac{r^2 \xi(r)}{\int_0^r x \xi(x) dx} = 0, \quad (9.7)$$

then  $\exists r_M^*$  such that,  $\forall r_M > r_M^*$ ,

$$E[\xi(\|L_i^{(2)} - l_e\|)] > E[\xi(\|L_i^{(1)} - l_e\|)], \forall l_e \in \partial B_{r_M}(0) \cap S_e. \quad (9.8)$$

*Proof:* Without loss of generality, suppose that  $l_e = (0, r_M)$ . This choice can be made because the uniform distribution imposed on  $L_i$  is invariant to rotations, which causes  $E[\xi(\|L_i - l_e\|)]$  to also be invariant to changes in  $l_e \in \partial B_{r_M}(0) \cap S_e$ , as shown in detail in Lemmas C.5 and C.4.

Consider first the case in which  $S_e$  is a regular polygon. As shown in the Appendix J,

$$E[\xi(\|L_i^{(1)} - (0, r_M)\|)] = \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} \xi(r) \cdot r dr d\omega, \quad (9.9)$$

$$E[\xi(\|L_i^{(2)} - (0, r_M)\|)] = \frac{1}{m(S_s^{(*2)})} \int_{-\pi}^{+\pi} \int_0^{h_n^{(2)}(\omega)} \xi(r) \cdot r dr d\omega, \quad (9.10)$$

where  $m(S_e)$  is the area of  $S_e$ ;  $h_n^{(1)}(\omega)$  determines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega) \in S_s^{(*1)}$ ; and  $h_n^{(2)}(\omega)$  determines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega) \in S_s^{(*2)}$ , where  $S_s^{(*j)}$  represents the deployment region when using option  $j$  translated by  $(0, -r_M)$  and rotated by  $\pi$ .

Decompose and bound  $E[\xi(\|L_i^{(2)} - (0, r_M)\|)]$  as follows

$$\begin{aligned} E[\xi(\|L_i^{(2)} - (0, r_M)\|)] &= \frac{1}{m(S_s^{(*2)})} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(2)}(\omega)} \xi(r) \cdot r dr d\omega \\ &\quad + \frac{2}{m(S_s^{(*2)})} \int_{\pi-\pi/n}^{3\pi/2} \int_0^{h_n^{(2)}(\omega)} \xi(r) \cdot r dr d\omega \end{aligned} \quad (9.11)$$

$$\geq \frac{1}{m(S_s^{(*2)})} \frac{m(S_e)}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} \xi(r) \cdot r dr d\omega \quad (9.12)$$

$$\begin{aligned} &\quad + \frac{2}{m(S_s^{(*2)})} \int_{\pi-\pi/n}^{3\pi/2} \int_0^{h_n^{(2)}(\omega)} \xi(r) \cdot r dr d\omega \\ &\geq \frac{1}{c^2} E[\xi(\|L_i^{(1)} - (0, r_M)\|)] \end{aligned} \quad (9.13)$$

$$+ \frac{1}{m(S_s^{(*2)})} \left( \pi + 2 \frac{\pi}{n} \right) \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) \cdot r dr, \quad (9.14)$$

where (9.11) is reached from  $h_n^{(2)}(\omega) = h_n^{(2)}(\pi - \omega)$  because of the symmetry of  $S_s^{(*2)}$  over the line  $x = 0$ ; (9.12) follows from the fact that  $h_n^{(2)}(\omega) > h_n^{(1)}(\omega)$  for all  $\omega \in [\pi/n, \pi - \pi/n]$  since  $S_s^{(*1)} \subset S_s^{(*2)}$ ; (9.13) follows from (9.9)

and the scaling property of the Lebesgue measure on  $\mathbb{R}^2$ ; and (9.14) follows because  $h_n^{(2)}(\omega) > (c-1)r_M \cos(\pi/n)$  because  $B_{(c-1)r_M \cos(\pi/n)}(0) \subset S_s^{(*2)}$ .

Decompose and bound  $E[\xi(\|L_i^{(1)} - (0, r_M)\|)]$  as follows

$$\begin{aligned} E[\xi(\|L_i^{(1)} - (0, r_M)\|)] &= \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} \xi(r) \cdot r \cdot \mathbf{1}\{r < (c-1)r_M \cos(\pi/n)\} dr d\omega \\ &\quad + \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} \xi(r) \cdot r \cdot \mathbf{1}\{r \geq (c-1)r_M \cos(\pi/n)\} dr d\omega \\ &\leq \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) \cdot r dr d\omega \end{aligned} \quad (9.15)$$

$$+ \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} \xi((c-1)r_M \cos(\pi/n)) \cdot r dr d\omega \quad (9.16)$$

$$\begin{aligned} &= \frac{1}{m(S_e)} \left( \pi - 2\frac{\pi}{n} \right) \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) \cdot r dr \\ &\quad + \xi((c-1)r_M \cos(\pi/n)) \cdot \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} r dr d\omega \end{aligned} \quad (9.17)$$

$$\leq \frac{1}{m(S_e)} \left( \pi + 2\frac{\pi}{n} \right) \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) \cdot r dr \quad (9.18)$$

$$+ \xi((c-1)r_M \cos(\pi/n)), \quad (9.19)$$

where (9.16) follows because  $\xi(r)$  is decreasing; (9.18) follows because  $\xi(r) \geq 0$ ; and (9.19) follows because the term multiplying  $1/m(S_e)$  in (9.17) is just the area of  $S_s^{(1)}$  in polar coordinates, which equals  $m(S_e)$ .

Use the bounds of (9.14) and (9.19) as follows: from (9.14),  $E[\xi(\|L_i^{(2)} - (0, r_M)\|)] - E[\xi(\|L_i^{(1)} - (0, r_M)\|)] > 0$  follows if

$$\frac{1}{m(S_s^{(*2)})} \left( \pi + 2\frac{\pi}{n} \right) \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) \cdot r dr + \frac{1-c^2}{c^2} E[\xi(\|L_i^{(1)} - (0, r_M)\|)] > 0. \quad (9.20)$$

Since  $c > 1$ ,  $1-c^2 < 0$ , and use the upper bound of (9.19) to conclude that (9.20) follows if

$$\begin{aligned} &\frac{1}{m(S_s^{(*2)})} \left( \pi + 2\frac{\pi}{n} \right) \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) r dr \\ &+ \frac{1-c^2}{c^2} \left[ \frac{1}{m(S_e)} \left( \pi + 2\frac{\pi}{n} \right) \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) r dr + \xi((c-1)r_M \cos(\pi/n)) \right] > 0. \end{aligned} \quad (9.21)$$

Using  $m(S_s^{(*2)}) = c^2 m(S_e)$ , rearrange (9.21) and obtain that (9.20) follows if

$$(2-c^2) \cdot \left( \pi + 2\frac{\pi}{n} \right) \cdot \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) r dr + (1-c^2) \cdot m(S_e) \cdot \xi((c-1)r_M \cos(\pi/n)) > 0. \quad (9.22)$$

The area of a polygon with  $n$  vertices and circumscribed by  $r_M$ ; i.e.,  $m(S_e)$ , equals  $r_M^2 n \sin(\pi/n) \cos(\pi/n)$  [34, p. 3] and (9.22) is equivalent to

$$(2-c^2) \cdot \left( \pi + 2\frac{\pi}{n} \right) \cdot \int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) r dr$$

$$+ (1 - c^2) \cdot \frac{n \sin(\pi/n) \cdot (c-1)^2 r_M^2 \cos^2(\pi/n)}{(c-1)^2 \cos(\pi/n)} \cdot \xi((c-1)r_M \cos(\pi/n)) > 0 \quad (9.23)$$

$$\Leftrightarrow (2 - c^2) \cdot \left( \pi + 2 \frac{\pi}{n} \right) + (1 - c^2) \cdot \frac{n \tan(\pi/n)}{(c-1)^2} \cdot \frac{(c-1)^2 r_M^2 \cos^2(\pi/n) \xi((c-1)r_M \cos(\pi/n))}{\int_0^{(c-1)r_M \cos(\pi/n)} \xi(r) r dr} > 0. \quad (9.24)$$

In order to see that (9.24) is satisfied by the conditions of this theorem, observe that the first term of is a positive constant since  $c < \sqrt{2}$ , and the second term is negative since  $c > 1$ ; however, the assumption (9.7) causes the second term to converge to 0 as  $r_M \rightarrow \infty$ .

The same steps are used to prove the  $S_e = B_{r_M}(0)$  case. In details, all steps in between the expressions (9.9) and (9.22) are valid once the terms  $(\pi/n)$  are replaced with 0. Using  $m(S_e) = \pi r_M^2$ , it is possible to reach that the modified (9.22) is equivalent to

$$(2 - c^2)\pi + (1 - c^2) \cdot \frac{\pi}{(c-1)^2} \cdot \frac{(c-1)^2 r_M^2 \xi((c-1)r_M)}{\int_0^{(c-1)r_M} \xi(r) r dr} > 0. \quad (9.25)$$

and the conclusion follow since the second term converges to 0 as  $r_M \rightarrow \infty$ . ■

It is possible to verify that the conditions of Proposition 9.1 are met when

- $\xi(r) = A_{\max} \exp\{-r^\kappa\}$  for any  $\kappa > 0$ ;
- $\xi(r) = A_{\max}/(1+r)^\kappa$  for any  $\kappa \geq 2$ ; <sup>4</sup> or
- $\xi(r)$  vanishes at some  $r = r_0$ .

The case  $\xi(r) = A_{\max}/(1+r)^\kappa$  for  $\kappa \in [1, 2)$  does not meet condition (9.7); however, Propositions K.1 and K.2 in the appendix conclude that there is a  $c$  and a large enough  $r_M$  such that the conclusion (9.8) is achieved for  $\kappa \in [1, 2)$  as well when  $S_e = B_{r_M}(0)$ .

## 9.4 Numerical Evaluations

Figure 9.2 presents numerical evaluations of  $ARE_{21}$  when  $S_e = B_{r_M}(0)$ ,  $S_e$  is an equilateral triangle circumscribed by  $B_{r_M}(0)$ , and  $S_e$  is a square circumscribed by  $B_{r_M}(0)$ . Formulas in Appendix J were used to generate this figure.

These figures confirm that, for large enough  $r_M$ , option 2 offers a performance gain over option 1 when  $\xi(r) = 1/(1+r)$ ,  $\xi(r) = 1/(1+r)^2$ , or  $\xi(r) = \exp\{-r\}$ .

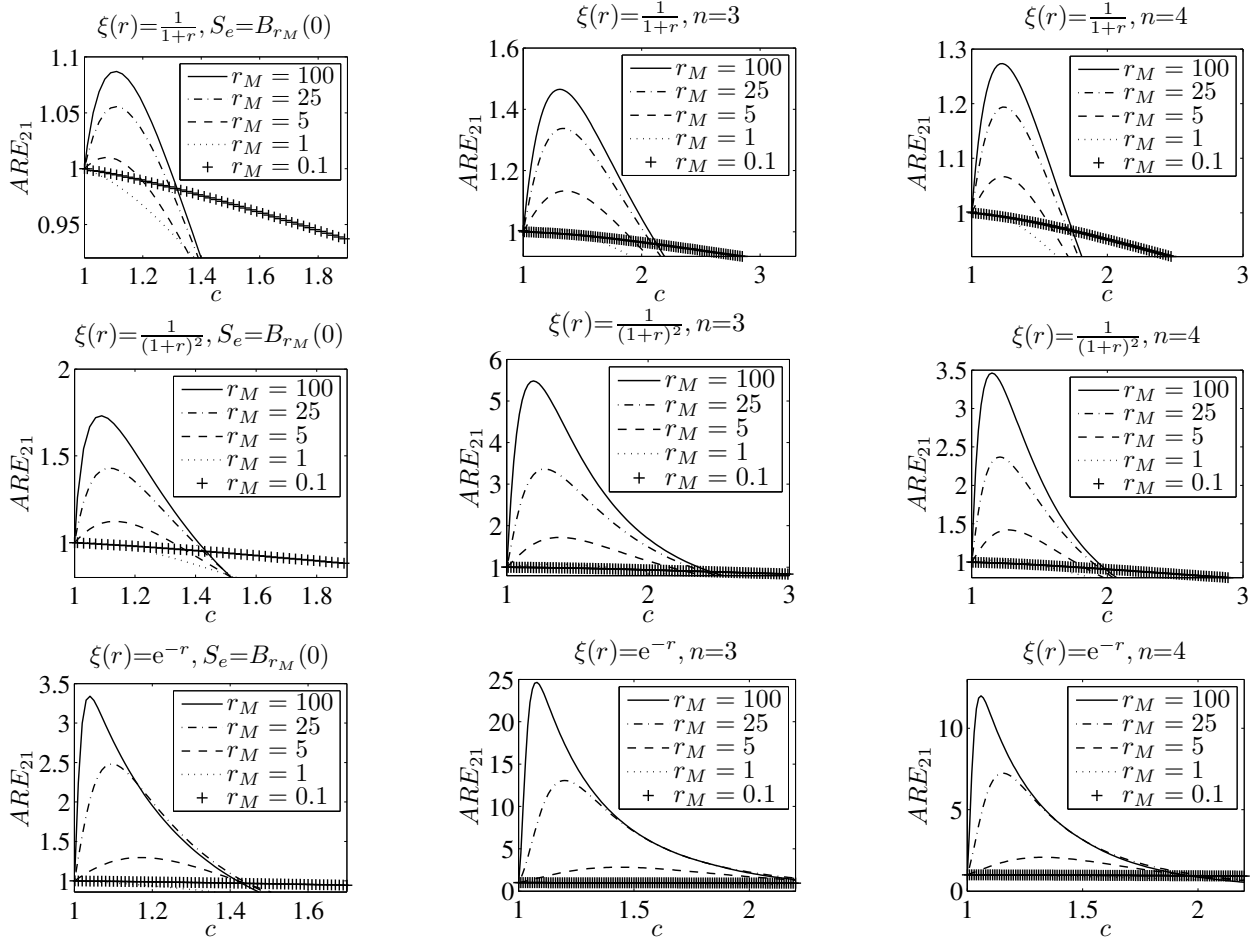
<sup>4</sup>To see why it satisfies for  $\kappa = 2$ , observe that  $\int r \xi(r) dr = \log(r+1) + 1/(r+1)$  and the ratio of (9.7) equals

$$\frac{x^2/(1+x)^2}{\log(x+1) + 1/(x+1) - 1} \quad (9.26)$$

which converges to 0 as  $x \rightarrow \infty$ . To see why it satisfies for  $\kappa > 2$ , observe that  $\int r \xi(r) dr = (1+r)^{(1-\kappa)}/(\kappa-1) - (1+r)^{(2-\kappa)}/(\kappa-2)$  and the ratio of (9.7) equals

$$\frac{x^{(2-\kappa)}/(1/x+1)}{x^{(1-\kappa)}/(1/x+1)/(\kappa-1) + x^{(2-\kappa)}/(1/x+1)/(\kappa-2) - (1/(\kappa-1) - 1/(\kappa-2))}$$

which converges to 0 as  $x \rightarrow \infty$ .



**Figure 9.2.** ARE<sub>21</sub> when  $\xi(r) = 1/(1+r)$  (top),  $\xi(r) = 1/(1+r)^2$  (middle), and  $\xi(r) = \exp(-r)$  (bottom) for various values of  $r_M$  when  $S_e = B_{r_M}(0)$  (left),  $S_e$  is a triangular region (middle), and  $S_e$  is a square region (right).

By comparing the potential performance gains offered by option 2 in Figure 9.2, one can infer that the potential performance gain increases significantly when  $S_e$  is a convex polygon. To understand this phenomenon, suppose for this paragraph that  $\xi(r) = \mathbf{1}\{r < r_0\}$  for some  $r_0 \in [0, r_M]$  and let  $S_n^{(j)} := \{(x, y) : \|(x, y) - l_e\| \leq r_0, (x, y) \in S_s^{(j)}\}$ ; i.e.,  $S_n^{(j)}$  is the set of points in  $S_s^{(j)}$  for which  $\xi(\|(x, y) - l_e\|) > 0$ , where the subscript  $n < \infty$  is used to denote the number of the vertices on  $S_e$  and  $n = \infty$  is used to denote the case in which  $S_e = B_{r_M}(0)$ . When using deployment option 1 and assuming  $l_e$  located at one of its vertices,  $m(S_3^{(1)}) < m(S_4^{(1)}) < m(S_n^{(1)}) < m(S_\infty^{(1)})$ , where  $m(S)$  represents Lebesgue measure and, therefore, the area of  $S$ . When using deployment option 2 with  $c = 1 + r_0/r_M$ ,  $S_n^{(2)}$  equals an open ball of radius  $r_0$ , which is contained in  $S_s^{(2)}$ . Thus,  $m(S_3^{(2)}) - m(S_3^{(1)}) > m(S_4^{(2)}) - m(S_4^{(1)}) > m(S_n^{(2)}) - m(S_n^{(1)}) > m(S_\infty^{(2)}) - m(S_\infty^{(1)})$ . Thus, when comparing the triangle, square, and ball cases, the performance gain offered by option 2 when  $S_s$  is an equilateral triangle is higher than the performance gain when  $S_s$  is a square, which is higher than the performance gain when  $S_s$  is a ball.



## 9.5 Upper Bound for the Performance Gain of Deployment Option 2

From the various situations presented in Figure 9.2, one may wonder if the performance gain of option 2 over option 1 always increases, for an appropriate  $c$ , as  $r_M$  increases towards infinity; however, as shown in this section, the performance gain of the deployment option 2 is bounded.

**Proposition 9.2.** *Let  $\xi$  be a nonincreasing and nonnegative function,  $L_i^{(1)}$  be uniformly distributed in  $S_s = S_e$ , and  $L_i^{(2)}$  be uniformly distributed in  $S_s = cS_e$  for any  $c > 1$ .*

When  $S_e = B_{r_M}(0)$ ,

$$ARE_{21} \leq 9 \cdot \left( 1 + \frac{\int_{r_M}^{c r_M} \xi(r) r dr}{\int_0^{r_M} \xi(r) r dr} \right)^2, \quad (9.27)$$

which converges to 9 as  $r_M \rightarrow \infty$ .

When  $S_e$  is a regular convex polygon with  $n$  vertices circumscribed by  $B_{r_M}(0)$ ,

$$ARE_{21} \leq \left( \frac{2n}{n-2} \right)^2 \cdot \left( 1 + \frac{\int_{r_M}^{c r_M} \xi(r) r dr}{\int_0^{r_M} \frac{2r_M \sin^2(\pi/n)}{2r_M \sin^2(\pi/n)} \xi(r) r dr} \right)^2, \quad (9.28)$$

which converges to  $(2n/(n-2))^2$  as  $r_M \rightarrow \infty$ .

*Proof:* To reach the result, an upper bound for  $E[\xi(\|L_i^{(2)} - (0, r_M)\|)]$  and a lower bound for  $E[\xi(\|L_i^{(1)} - (0, r_M)\|)]$  are derived.

For the upper bound, observe that  $E[\xi(\|L_i^{(2)} - (0, r_M)\|)] \leq E[\xi(\|L_i^{(2)} - (0, 0)\|)]$ , which follows if  $A_i = \xi(\|L_i^{(2)} - L_e\|)$  and  $h(a) = a$  in the steps between (7.51) and (7.64) in the proof of Proposition 7.7, and, using polar coordinates, write

$$E[\xi(\|L_i^{(2)} - (0, r_M)\|)] \leq \frac{2\pi}{c^2 m(S_e)} \int_0^{c r_M} \xi(r) r dr \leq \frac{2\pi}{m(S_e)} \int_0^{c r_M} \xi(r) r dr. \quad (9.29)$$

For the lower bound, consider first that  $S_e = B_{r_M}(0)$  and, as shown in the proof of Proposition 8.5, relation (8.29),

$$E[\xi(\|L_i^{(1)} - (0, r_M)\|)] \geq \frac{1}{3} \frac{2\pi}{m(S_e)} \int_0^{r_M} \xi(r) r dr, \quad (9.30)$$

and, using (9.29) and (9.30), reach that

$$\frac{E[\xi(\|L_i^{(2)} - (0, r_M)\|)]}{E[\xi(\|L_i^{(1)} - (0, r_M)\|)]} \leq 3 \cdot \left( 1 + \frac{\int_{r_M}^{c r_M} \xi(r) r dr}{\int_0^{r_M} \xi(r) r dr} \right), \quad (9.31)$$

and the bound (9.27) follows.

The bound (9.28) follows similarly. As shown in Proposition 8.4, relation (8.21),

$$E[\xi(\|L_i^{(1)} - (0, r_M)\|)] \geq \frac{\pi - 2\pi/n}{m(S_e)} \int_0^{2r_M \sin^2(\pi/n)} \xi(r) r dr, \quad (9.32)$$

and it follows that

$$\frac{E[\xi(\|L_i^{(2)} - (0, r_M)\|)]}{E[\xi(\|L_i^{(1)} - (0, r_M)\|)]} \leq \frac{2n}{n-2} \cdot \left(1 + \frac{\int_{2r_M \sin^2(\pi/n)}^{c r_M} \xi(r) r dr}{\int_0^{2r_M \sin^2(\pi/n)} \xi(r) r dr}\right). \quad (9.33)$$

■

Since the upper bound for the  $ARE_{21}$  converges to 36 when  $n = 3$  and to 16 when  $n = 4$ , it is possible to see that the numerical evaluations present in Figure 9.2 are in agreement with the bounds of Proposition 9.2.

## 9.6 The Impact of the Deployment Option 2 in the Performance Under other Distributions for the Emitter Location

In the previous sections, it was shown that there are conditions under which the deployment option 2 improves the detection performance when the emitter location  $L_e$  is distributed according to a least favorable distribution for the system; however, it does not necessarily improve, and may in fact degrade, the detection performance when  $L_e$  is distributed according to other distributions.

For instance, assuming the conditions of any one of the Propositions 8.2, 8.3, and 8.8 and letting  $S_e = B_{r_M}(0)$ , if  $L_e$  is distributed according to the most favorable distribution for the system; i.e.,  $P[L_e = (0, 0)] = 1$ , then the deployment option 2 always degrades the performance of the system. To see this, use the formulas of Appendix J to see that the ARE of the system using option 2 with respect to option 1 is given by

$$\frac{E[\xi(\|L_i^{(2)} - (0, 0)\|)]}{E[\xi(\|L_i^{(1)} - (0, 0)\|)]} = \frac{1}{c^2} \frac{\int_0^{c r_M} \xi(r) \cdot r dr}{\int_0^{r_M} \xi(r) \cdot r dr} = \frac{1}{c^2} \left[1 + \frac{\int_{r_M}^{c r_M} \xi(r) \cdot r dr}{\int_0^{r_M} \xi(r) \cdot r dr}\right] \leq 1, \quad (9.34)$$

where the assumption that  $\xi(r)$  is nonincreasing was used to reach the inequality in (9.34). This conclusion can also be reached in the example of Section 9.1, since  $m(S_s^{(2)} \cap B_{r_0}^\circ(0)) = m(S_s^{(1)} \cap B_{r_0}^\circ(0))$  if  $r_0 \leq r_M$  and  $m(S_s^{(2)}) > m(S_s^{(1)})$ , which means that (9.2) is never satisfied for this case.

This means that, although the deployment option 2 may provide a better detection performance than option 1 when  $L_e$  is distributed according to the least favorable distribution for a system, it may cause worse detection performance than option 1 for the actual, unknown, distribution for  $L_e$ .

In spite of this drawback, it is argued here that, when the deployment option 2 offers a better detection performance than option 1 for the least favorable distribution, the deployment option 2 should be preferred over option 1 for the following reasons:

- The deployment option 2 provides a higher minimum detection performance over all possible distributions for  $L_e$ , which corresponds to a lower number of sensors to satisfy a detection requirement; and,
- Even though the detection performance when using deployment option 2 may be worse than option 1 when  $L_e$  has a different distribution, it is guaranteed that such a performance will still be above the minimum performance

requirement if option 2 satisfies the minimum performance requirement under the least favorable distribution. In other words, if the goal is to satisfy a prescribed performance, then the degradation in performance under a particular distribution for  $L_e$  is not a problem as long as it stays above the minimum performance requirement.

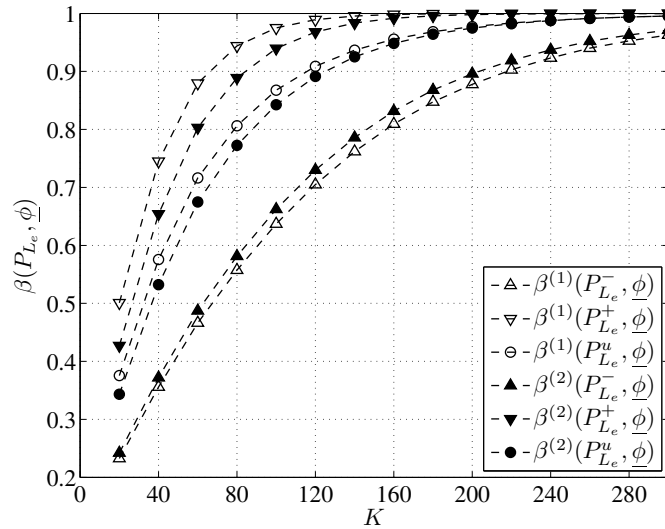
The above arguments are illustrated in the next example.

### 9.6.1 Example

In the example of Section 8.1, it was assumed that the deployment region  $S_s = cS_e$  with  $c = 1$ ; i.e., deployment option 1 was considered. In this section, this example will be changed to evaluate the performance under the deployment option 2. Thus, assume that  $S_s = cS_e$  for  $c = 1.1$  and assume all other conditions of the example.

Using numerical integration of the equations given in Appendix H, it is possible to evaluate the probabilities of detection under a least favorable distribution  $P_{L_e}^-$  for  $\underline{\phi}$ , under the most favorable distribution  $P_{L_e}^+$  for  $\underline{\phi}$ , and under the uniform distribution  $P_{L_e}^u$  for various values of  $K$ . Taking  $\alpha_{\max} = 0.05$ ,  $\lambda_W = 1$ ,  $t_s = \lambda_W$ ,  $r_M = 15$ , and  $A_{\max} = 40$ , the resulting probabilities of detection under the deployment option 1 ( $\beta^{(1)}(P_{L_e}, \underline{\phi})$ ) and under the deployment option 2 ( $\beta^{(2)}(P_{L_e}, \underline{\phi})$ ) are shown in Figure 9.3.

From Figure 9.3, it is possible to see that  $\beta^{(1)}(P_{L_e}^+, \underline{\phi}) > \beta^{(2)}(P_{L_e}^+, \underline{\phi})$  and  $\beta^{(1)}(P_{L_e}^u, \underline{\phi}) > \beta^{(2)}(P_{L_e}^u, \underline{\phi})$ ; i.e., the probability of detection when  $L_e$  is distributed according to either the most favorable distribution for  $\underline{\phi}$  or the uniform distribution decreases when deployment option 2 is used instead of deployment option 1 for every  $K$ ; however, the



**Figure 9.3.** Probability of detection of the distributed detection system defined by (8.5) and (8.6) with the distribution of  $Y_0$  and  $k$  defined such that the probability of false alarm equals  $\alpha_{\max} = 0.05$  when  $L_e$  is distributed according to  $P_{L_e}^-$  that places the emitter in the boundary of a circular region of interest with probability 1,  $P_{L_e}^+$  that satisfies  $P[L_e = (0, 0)] = 1$ , and the uniform distribution  $P_{L_e}^u$ ; and when  $\{L_i\}_{i=1}^K$  is uniformly distributed in  $S_s = cS_e$  for either  $c = 1$  (deployment option 1) or  $c = 1.1$  (deployment option 2) for various values for the number of sensors  $K$ .

probability of detection when  $L_e$  is distributed according to the least favorable distribution for  $\underline{\phi}$  increases when deployment option 2 is used instead of deployment option 1.

If a system designer imposes a minimum probability of detection  $\beta_{\min} = 0.95$ , then the resulting design under deployment option 2 requires  $K^{(\Phi, P_{L_e}^-)} \approx 260$  to satisfy the requirement, which is less than the approximately 280 sensors required under the deployment option 1, and the requirement is still satisfied for any other distribution for  $L_e$ .

## 9.7 Summary

In this section, the ARE metric was used to compare the deployment option in which the sensor deployment region is equal to the region of interest (option 1) against the deployment option in which sensors are deployed in an enlarged deployment region (option 2). Assuming the Poisson and Gaussian cases, and assuming the same conditions used in various propositions of Chapter 8, the efficacy of many sensor detection systems of interest depends on the distribution of sensor locations  $L_i$  through the factor  $E[\xi(\|L_i - (0, r_M)\|)]$ . With such a result, the ARE metric is the ratio between the efficacies of the two deployment options under the least favorable distribution for the detection system, as displayed in (9.6).

Proposition 9.1 showed that there are conditions under which option 2 provides a better asymptotic detection performance than option 1. Such conditions are met for many cases of interest, and the propositions in the Appendix K cover additional cases of interest. Numerical evaluations illustrate that the performance gains of option 2 over option 1 can be significant, particularly when  $S_e$  is an equilateral triangle or a square. Such evaluations illustrate that the performance gain increases with  $r_M$ ; however, Proposition 9.2 shows that there is an upper bound for the performance gain of option 2 over option 1.

It was observed that the deployment option 2 has a drawback in that it may cause a better detection performance than option 1 for the least favorable distribution while causing a worse detection performance than option 1 for the actual, unknown, distribution for  $L_e$ . Nevertheless, it is argued that the deployment option 2 should be preferred over option 1 because it provides a higher guaranteed minimum detection performance for all possible distributions for  $L_e$ .

In conclusion, when the conditions of Proposition 9.1 are satisfied, a system designer can adopt the deployment option 2 to reduce the number of additional sensors required by the adoption of the least favorable distribution for the emitter location.

## Chapter 10

### Summary of Contributions

1. Propositions 6.2, 6.3, 6.4, and 6.5 identify sufficient conditions under which the measurements are conditionally i.i.d. given  $\theta$  in several scenarios. Such sufficient conditions include having: sensor locations uniformly distributed in the deployment region, and regions of interest that are in the form of a circle, sphere, or regular convex polygon.
2. The theory of least favorable distributions is applied to the distribution of the emitter location. Although the use of least favorable distributions is not new, to the best of the author's knowledge, this is the first time that the theory of least favorable distributions is applied to the emitter location distribution.
3. Inadequacies in the concept of the least favorable distribution are indicated and the concept is augmented with the definition of the concepts of a least favorable distribution for a class of systems and a least favorable distribution for a system.
4. A least favorable emitter location distribution for detection systems of interest, including the maximin system, is identified. Under conditions that include having sensor locations uniformly distributed and regions of interest in the form of a circle or a regular convex polygon, a least favorable emitter location distribution for several systems of interest places the emitter on the boundary of the circular region of interest or at the vertices of the polygonal region of interest. It is shown that this distribution is not always least favorable.
5. It is shown that the least favorable emitter location distribution for detection systems of interest, including the maximin system, is one of the distributions that cause the measurements to become conditionally i.i.d. This result indicates that, if the conditions of Propositions 7.7 or 7.8 are satisfied and a system designer adopts the identified least favorable emitter location distribution, then the designer will not only have a simple hypothesis and ensure a detection performance, but also avoid the problem of conditional dependence among measurements.

6. The concept of a most favorable emitter location distribution for a detection system and the concept of a max-max system are defined. A most favorable emitter location distribution for detection systems of interest, including the maximax system, is identified. Under conditions that include having sensor locations uniformly distributed and regions of interest in the form of a circle or a regular convex polygon, a most favorable emitter location distribution for several systems of interest places the emitter at the center of the region of interest.
7. It is shown that a most favorable emitter location distribution for a system can be used to determine how conservative a design under a least favorable emitter location distribution is in the asymptotic regime. For such a determination, it is proposed that the system designer use the ARE metric of the design under a most favorable emitter location distribution with respect to the design under a least favorable emitter location distribution.

When the measurement noise is Gaussian or Poisson distributed, formulas are provided to allow the computation of the ARE metric among various combinations of systems, including the maximin, the maximax, a practical centralized detection system  $\Phi(\Sigma)$  in which the fusion function decides based upon the sum of the measurements, and a practical distributed detection system  $\Phi(\Sigma,dds)$  in which the fusion function decides based upon the sum of binary sensor functions that compare the measurement against a threshold.

It is shown that the ARE of the maximax centralized system with respect to the maximin centralized system, the ARE of  $\Phi(\Sigma)$  under a most favorable emitter location distribution with respect to  $\Phi(\Sigma)$  under a least favorable emitter location distribution, and the ARE of  $\Phi(\Sigma,dds)$  under a most favorable emitter location distribution with respect to  $\Phi(\Sigma,dds)$  under a least favorable emitter location distribution are all given by the same expression:

$$\left( \frac{E[\xi(\|L_i - (0,0)\|)]}{E[\xi(\|L_i - (0,r_M)\|)]} \right)^2, \quad (10.1)$$

which is amenable for numerical computation. Analytical bounds are provided for guidance. Using these bounds, it is shown that a design under a least favorable emitter location distribution in the asymptotic regime would need at most 9 times more sensors than a design under the most favorable emitter location distribution when the region of interest is circular. When the region of interest is an equilateral triangle or a square, this number increases to 36 or 16 respectively. If the region of interest is large enough, it is shown that the upper bound for the circular region of interest reduces to 4.

These results indicate to the system designer how many more sensors need to be deployed to compensate for the lack of knowledge about the emitter location distribution.

8. Formulas for the ARE metric among various systems of interest under a common least favorable emitter location distribution and under Gaussian or Poisson measurement noise are provided:

- It is shown that the ARE of the maximin centralized detection system with respect to  $\Phi(\Sigma)$ , both under a least favorable emitter location distribution, equals 1. This result shows that there is no loss in performance when the system designer replaces the maximin centralized detection system with such a practical centralized detection system.<sup>1</sup>
  - A formula is provided to compute the ARE of  $\Phi(\Sigma)$  with respect to  $\Phi(\Sigma,dds)$ , both under a least favorable emitter location distribution. It is shown that the ARE metric is a function of the statistics of the measurement noise distribution. When the measurement noise is Gaussian distributed, it is shown that the ARE is minimized when common sensor thresholds equal 0 and, in such a case, the ARE metric equals  $\pi/2$ . Such a value for an ARE metric was reported by previous authors in the simpler case of deterministic known signals. This result extends the previous result to the case in which the signal is randomly distributed with a distribution that depends on the distance between sensors and emitter. This result helps a system designer evaluate how many more sensors are required when adopting a distributed design.
  - It is shown that the ARE of the maximax detection system with respect to  $\Phi(\Sigma,dds)$  is given by the product of two factors, the first one related to a least favorable emitter location distribution and given by (10.1), and the second one related to the type of detection system: centralized or distributed. This result helps a system designer evaluate the contribution of each factor in the final number of additional sensors required to compensate for the use of a practical system and a least favorable emitter location distribution.
9. A method to compare sensor deployment options under a least favorable emitter location distribution is provided, and it is shown that there are sufficient conditions under which deploying sensors in a larger region than the region of interest provides improved asymptotic detection performance. This result is counterintuitive because it proposes that sensors be deployed in locations known to not contain the signal emitter. Numerical evaluations and analysis are provided, showing that the performance gain can be significant, particularly when the region of interest is a regular convex polygon with a small number of vertices. This result indicates that the adoption of larger sensor deployment regions can be used to reduce the additional number of sensors required by the adoption of a least favorable emitter location distribution.

The conclusions that can be inferred from the above contributions are presented in the next chapter.

---

<sup>1</sup>This result would be trivial if the measurements were distributed according to the Gaussian or the Poisson distributions; however, because of the random locations of emitter and sensors, the measurements are distributed according to a mixture of Gaussian or Poisson distributions and the sum of measurements is not a sufficient statistic for decision.

## Chapter 11

### Conclusions and Future Research

The first conclusion of this dissertation is that there are certain sets of conditions under which the problem of conditionally dependent measurements disappears and the measurements become conditionally i.i.d. given the hypotheses. One such set of conditions includes having the distribution of the emitter location place the emitter in a subset of a circular or polygonal region of interest with probability one. Another set of conditions includes having the amplitude of the emitted signal eventually decay to zero; and another set of conditions includes having the region of interest be a sphere. These sets of conditions contradict the commonly held belief that, if sensors measure a signal that decays with the distance, then measurements are conditionally dependent. This dissertation shows that a system designer does not need to worry about intractable models caused by conditional dependence if any of the identified sets of conditions holds.

The second and perhaps most important conclusion of this dissertation is that there are conditions under which a least favorable emitter location distribution for several systems of interest is one of the distributions that causes the measurements to become conditionally i.i.d. This result indicates that, if the conditions are satisfied and the system designer adopts the identified least favorable emitter location distribution, then the designer will not only ensure a detection performance and avoid the difficulties of a composite hypothesis, but will also avoid the difficulties associated with conditionally dependent measurements. This conclusion offers a better justification for assuming conditionally i.i.d. measurements than the commonly used justifications. As described in Chapter 4, many authors argued that the assumption of conditionally i.i.d. measurements is justified by convenience or by lack of information about the correlation parameters.

The third conclusion of this dissertation is that it is possible to evaluate how conservative a design under a least favorable emitter location distribution for a system is by using the concept of a most favorable emitter location distribution. When the number of sensors is large and the measurement noise is either Gaussian or Poisson distributed, such an evaluation can be performed for many of the systems of interest through the computation of a simple expression involving the amplitude function  $\xi$  and the dimensions of the region of interest. Upper and lower bounds for this expression were provided for the circular and regular convex polygonal regions of interest. If the region of interest is circular and large enough, then it is shown that such an expression is close to 4, meaning that a system designed under



the least favorable emitter location distribution would require at most 4 times as many sensors than the system designed under the actual emitter location distribution to reach the same asymptotic detection performance. This multiplication factor increases to 16 and 36 respectively when the region of interest is square or triangular.

The fourth conclusion of this dissertation is that, once a single distribution for the emitter location is found and the measurements are conditionally i.i.d., then a system designer can use the tools available for the conditionally-independent and simple-hypothesis setting. This conclusion was illustrated in the computation of the ARE metric among several systems of interest under a common least favorable emitter location distribution.

The fifth conclusion of this dissertation is that there are situations in which deploying sensors in a region larger than the region of interest can improve the asymptotic detection performance. When the conditions are satisfied, larger deployment regions can be used by a system designer to reduce the number of sensors required to ensure a given detection performance under the least favorable emitter location distribution.

All of the above conclusions support this dissertation thesis: if the conditions of Propositions 7.7 or 7.8 are satisfied, then the system designer can assume the emitter location to be distributed according to a least favorable distribution for the system, which will not only ensure a detection performance and avoid dealing with composite hypotheses, but also make the measurements conditionally i.i.d. Tools available for conditionally i.i.d. measurements can be used; for instance, the large sample method of ARE can be used to compare candidate designs and evaluate how conservative the design based on a least favorable emitter location distribution is.

## 11.1 Future Research

A first avenue for future research involves the extension of the main results to cases not considered in this dissertation.

Most of the results presented in this dissertation assume that the communication subsystem can offer a dedicated and error-free channel to each of the sensors. Since many sensor detection systems being envisioned by other authors consider non-dedicated or error-prone communication channels, as described in Chapter 3, it would be useful to extend the results of Chapter 7 to show that the least favorable emitter location distribution identified for the dedicated and error-free communication subsystem is still least favorable when considering either a communication subsystem offering dedicated but error-prone channels, or a communication subsystem offering multiple access channels as described in Section 3.1.5. The author expects that the results obtained in this dissertation will still hold even under such communication subsystems.

The results presented in this dissertation may also be extended to cover sensor detection systems using censoring. As described in Section 3.1.4.1, censoring of sensor transmissions is an effective technique for reducing energy consumption in sensors. It would be useful to extend the propositions of Chapter 7 to show that the least favorable emitter

location distribution identified for the dedicated and error-free communication subsystem is still least favorable when considering sensor detection systems with censoring.

The asymptotic results presented in Chapters 8 and 9 considered either the Gaussian or the Poisson cases, as defined in Section 8.2. Other signal and noise models may satisfy the conditions of Proposition I.3 and the conclusions of Chapters 8 and 9 would hold for such models as well.

All the results of this dissertation assumed a single emitter within the region of interest. As mentioned in Appendix A, considering a single emitter is a worst case assumption in applications such as radiation detection; however, the single emitter assumption is not necessarily a worst case assumption in other applications, such as when detecting radio transmissions.

The results presented in this dissertation may also be extended to cover multiple types of sensor measurements. In this dissertation, it was assumed that sensors collect measurements following a single signal model. Some applications involve sensors collecting different measurements, each one following a different signal model; for instance, in the application of detecting forest fires, it is common to use sensors collecting air temperature and humidity [57, 103, 104, 140].

This dissertation assumed that the hypothesis  $H_1$  was composite only because of the lack of knowledge about the emitter location distribution and  $H_1$  became simple with the adoption of the least favorable emitter location distribution. However, even when a distribution for the emitter location is adopted,  $H_1$  may still be composite if the signal level is considered unknown as well. In this case, it would be useful to extend the results of this dissertation to the locally optimum case.

This dissertation also assumed that the fusion center decides upon a single set of measurements from the sensors. Sequential or change-point detection methods can also be used and it would be helpful to determine whether the least favorable emitter location distribution identified in this dissertation would still be least favorable when using these methods.

A second avenue for future research involves the use of ARE theory and the results of Chapters 8 and 9 to compare further sensor deployment options.

Assuming that a system designer partitions a large region into several regions of interest, each one being monitored by a single sensor detection system, it would be useful to determine the best shapes to use in partitioning the large region. The theory and methods presented in Chapters 8 and 9 could be used to compare two candidate shapes for the region of interest.

This dissertation showed that it is possible to improve detection performance by using a sensor deployment region that is different from the region of interest; however, this result requires that the deployment region be a scaled version of the region of interest; i.e., both have the same shape. Could the performance improve if the sensor deployment region had a different shape from the region of interest?

A third avenue for future research involves the consideration of different distributions for the sensor locations.

The emitter location distribution identified in Chapter 7 is a least favorable distribution for a system assuming that the sensor locations are uniformly distributed in the deployment region. If a different distribution for the sensor locations were assumed, such a distribution for the emitter location might no longer be least favorable. From this observation, a system designer may wonder: which sensor location distribution would provide the highest detection performance when adopting its corresponding least favorable emitter location distribution?

## Appendix A: Conditions under which Considering a Single Emitter is the Worst Case Scenario

**Proposition A.1.** Suppose a detection system  $(\{\phi_i\}_{i=0}^K, P_Y)$  from either the class  $\mathcal{D}_c$  or the class  $\mathcal{D}_d$  of detection systems defined in Section 7.3. Assume that the number of sensors  $K$ , the fusion function  $\phi_0$ , the various sensor functions  $\{\phi_i\}$ , and the distribution  $P_Y$  of the randomization random variable  $Y$  were designed to satisfy a required probability of detection for a maximum probability of false alarm while assuming that each sensor measurement is given by

$$Z_i = \mathbf{1}\{\theta > 0\} \cdot A_i + W_i; \quad (\text{A.1})$$

where  $\theta \in \{0, \theta_1\}$ ,  $W_i$  represents the measurement noise, and  $A_i$  represents the signal received from a single emitter in the region of interest.

If  $J > 1$  emitters exist in the region of interest when  $\theta = \theta_1$  and the measurement  $Z_i$  at each sensor is instead given by

$$Z_i = \mathbf{1}\{\theta > 0\} \cdot h(A_i^{(1)}, \dots, A_i^{(J)}) + W_i, \quad (\text{A.2})$$

where  $A_i^{(j)}$  is the signal random variable received at sensor  $i$  from emitter  $j$ ; and  $h : \mathbb{R}^J \rightarrow \mathbb{R}$  satisfies<sup>1</sup>

$$\forall a^{(1)}, \dots, a^{(J)} \in \mathbb{R}, \forall j \in \{1, \dots, J\}, a^{(j)} \leq h(a^{(1)}, \dots, a^{(J)}); \quad (\text{A.3})$$

then the detection system  $(\{\phi_i\}_{i=0}^K, P_Y)$  still satisfies the required probability of detection for a maximum probability of false alarm.

*Proof:* Observe that the probability of false alarm is the same in both cases since the number of emitters does not affect the measurements at  $\theta = 0$ ; thus, it is enough to prove that the probability of detection increases under the measurement model of (A.2).

Suppose any detection system in  $\mathcal{D}_d$ , and write,  $\forall a_i^{(1)}, \dots, a_i^{(J)}, \forall w_i$ , and  $\forall y$ ,

$$a_i^{(1)} + w_i \leq h(a^{(1)}, \dots, a^{(J)}) + w_i$$

$$g(T_i(a_i^{(1)} + w_i), y) \leq g(T_i(h(a^{(1)}, \dots, a^{(J)}) + w_i), y) \quad (\text{A.4})$$

$$T_{0,i}(g(T_i(a_i^{(1)} + w_i), y)) \leq T_{0,i}(g(T_i(h(a^{(1)}, \dots, a^{(J)}) + w_i), y)) \quad (\text{A.5})$$

$$\prod_{i=1}^K T_{0,i}(g(T_i(a_i^{(1)} + w_i), y)) \leq \prod_{i=1}^K T_{0,i}(g(T_i(h(a^{(1)}, \dots, a^{(J)}) + w_i), y)), \quad (\text{A.6})$$

where (A.4) follows because  $T_i$  and  $g$  are nondecreasing, (A.5) follows because  $T_{0,i}$  is also nondecreasing, and (A.6) follows because  $T_{0,i}$  is nonnegative. Thus, for any set of realizations  $\{\{a_i^{(j)}\}_{j=1}^J\}_{i=1}^K$  of  $\{\{A_i^{(j)}\}_{j=1}^J\}_{i=1}^K$ , any set of

<sup>1</sup> An important case is  $h(a^{(1)}, \dots, a^{(J)}) = \sum_{j=1}^J a^{(j)}$ , which occurs when sensors count the emissions from multiple ionizing radioactive materials [20, 98, 106].

realizations  $\{w_i\}_{i=1}^K$  of  $\{W_i\}_{i=1}^K$ , and any realization  $y$  of  $Y$ ,

$$\prod_{i=1}^K T_{0,i}(g(T_i(a_i^{(1)} + w_i), y)) \in I_{0,y} \Rightarrow \prod_{i=1}^K T_{0,i}(g(T_i(h(a^{(1)}, \dots, a^{(J)}) + w_i), y)) \in I_{0,y}, \quad (\text{A.7})$$

since  $I_{0,y}$  is either  $(t_{0,y}, \infty]$  or  $[t_{0,y}, \infty]$ ; and (A.7) means that the probability of detection increases when  $J > 1$  emitters are present.

If the detection system is in  $\mathcal{D}_c$ , replace  $g(T_i(a_i^{(1)} + w_i), y)$  with  $a_i^{(1)} + w_i$  and replace  $g(T_i(h(a^{(1)}, \dots, a^{(J)}) + w_i), y)$  with  $h(a^{(1)}, \dots, a^{(J)}) + w_i$  in all the above steps. ■

## Appendix B: Scenarios in which Sensor Locations can be Considered Random Variables

1. Sensors may be mobile, which is the case when considering systems to detect transmissions of a primary user in a cooperative spectrum-sensing system [120, 126, 139] or when the sensors are carried by a person, animal, or vehicle [100].
2. Deployment of sensors in predetermined locations may not be possible [2]. For instance, the sensors may be released from an aircraft or contained within projectiles that scatter the sensors in a region [6, 40, 57, 58, 61, 83, 125].
3. While certain sensor detection systems are designed to operate for long periods of time, which justify careful consideration for the sensor placement, other sensor detection systems are designed to operate for a short period of time and to be deployed immediately before its operation [33, 125]; for instance, fire fighters may deploy a sensor detection system within a building in flames as they start fighting the fire [125]. In such situations, sensors are more likely to be randomly deployed.
4. It may be necessary to design the detection system *before* the sensor locations are defined or the deployment of sensors may not be under the control of the system designer [108]. In these cases, there is uncertainty about the sensor locations.
5. Even if the sensor locations are deterministic quantities, the set of sensors that transmit measurements to the fusion center in the decision interval may be random due to, for instance, censoring algorithms [17, 18].

Although sensors could be equipped with global positioning system (GPS) units and transmit its location to the fusion center during initialization or together with its measurement, equipping all sensors with GPS units may not be viable, particularly when sensors have a small form factor, when deploying a large number of sensors, or when sensors cannot receive the GPS signals [2, 43]. Furthermore, the transmission of locations to the fusion center is not possible if such locations need to be transmitted along with every measurement, which would be necessary if sensors are mobile; and the communication subsystem does not offer  $K$  dedicated communication channels, as envisioned by some authors [52, 74, 78, 81].

Moreover, even if the sensors' locations were readily available at the fusion center, they are not enough to determine the precise correlation among measurements and build a better performing fusion function.

## Appendix C: Supporting Lemmas for Chapter 6

**Lemma C.1.** *Adopt Assumptions II, III, IV, V, and VI of Chapter 5. If the distribution of  $L_e$  satisfies*

$$P[L_e \in S_e^*] = 1 \quad (\text{C.1})$$

*for some  $S_e^* \subseteq S_e$  that satisfies*

$$\forall i, \forall l_{e_1}, l_{e_2} \in S_e^*, \forall a, P[A_i \leq a | \theta = \theta_1, L_e = l_{e_1}] = P[A_i \leq a | \theta = \theta_1, L_e = l_{e_2}] =: v_i(a); \quad (\text{C.2})$$

*i.e.;  $P[A_i \leq a | \theta = \theta_1, L_e = l_e]$  does not change as  $l_e$  varies over  $S_e^*$  for all  $i$  and all  $a$ ; then*

1.  $\{A_i\}_{i=1}^K$  and  $L_e$  are conditionally independent given  $\theta = \theta_1$ ;
2. If  $\{L_i\}_{i=1}^K$  are i.i.d. random variables, then  $\{A_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta = \theta_1$ .

*Proof:* Observe first that, under the conditions of this proposition,  $P[A_i \leq a_i | \theta = \theta_1] = v_i(a_i)$ . To see this, let  $P_{L_e}$  be the probability measure of  $L_e$  and write

$$\begin{aligned} P[A_i \leq a_i | \theta = \theta_1] &= \int_{S_e^*} P[A_i \leq a_i | \theta = \theta_1, L_e = l_e] dP_{L_e}(l_e) + \\ &\quad \int_{(S_e^*)^c} P[A_i \leq a_i | \theta = \theta_1, L_e = l_e] dP_{L_e}(l_e) \end{aligned} \quad (\text{C.3})$$

$$= \int_{S_e^*} v_i(a_i) dP_{L_e}(l_e) = v_i(a_i) \cdot P[L_e \in S_e^*] = v_i(a_i), \quad (\text{C.4})$$

and note that the second term in the addition of (C.3) equals zero because of (C.1), and (C.4) follows from (C.2) and (C.1).

To reach the first conclusion, pick any  $\{a_i\}_{i=1}^K$  and any measurable set  $\mathcal{B} \subset \mathbb{R}^\delta$  and write

$$\begin{aligned} P \left[ \bigcap_{i=1}^K \{A_i \leq a_i\}, L_e \in \mathcal{B} \middle| \theta = \theta_1 \right] &= \int_{\mathcal{B} \cap S_e^*} P \left[ \bigcap_{i=1}^K A_i \leq a_i \middle| \theta = \theta_1, L_e = l_e \right] dP_{L_e}(l_e) + \\ &\quad \int_{\mathcal{B} \cap (S_e^*)^c} P \left[ \bigcap_{i=1}^K A_i \leq a_i \middle| \theta = \theta_1, L_e = l_e \right] dP_{L_e}(l_e) \\ &= \int_{\mathcal{B} \cap S_e^*} P \left[ \bigcap_{i=1}^K A_i \leq a_i \middle| \theta = \theta_1, L_e = l_e \right] dP_{L_e}(l_e) \end{aligned} \quad (\text{C.5})$$

$$= \int_{\mathcal{B} \cap S_e^*} \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1, L_e = l_e] dP_{L_e}(l_e) \quad (\text{C.6})$$

$$= \int_{\mathcal{B} \cap S_e^*} \prod_{i=1}^K v_i(a_i) dP_{L_e}(l_e) = \prod_{i=1}^K v_i(a_i) P[L_e \in \mathcal{B} \cap S_e^*] \quad (\text{C.7})$$

$$= \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1] \cdot P[L_e \in \mathcal{B} \cap S_e^*] \quad (\text{C.8})$$

$$= \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1] \cdot (P[L_e \in \mathcal{B} \cap S_e^*] + P[L_e \in \mathcal{B} \cap (S_e^*)^c]) \quad (\text{C.9})$$

$$= \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1] \cdot P[L_e \in \mathcal{B}] = \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1] \cdot P[L_e \in \mathcal{B} | \theta = \theta_1] \quad (\text{C.10})$$

where (C.5) follows from (C.1) because the integral of any measurable function over a set of measure zero is zero, (C.6) follows from Lemma C.2, (C.7) follows from (C.2), (C.8) follows from (C.4), (C.9) follows because  $P[L_e \in \mathcal{B} \cap (S_e^*)^c] = 0$ , and (C.10) follows because  $L_e$  and  $\theta$  are independent.

Since the first conclusion established that  $\{A_i\}_{i=1}^K$  are conditionally independent given  $\theta = \theta_1$ , the second conclusion follows once it is proven that  $\{A_i\}_{i=1}^K$  are identically distributed given  $\theta = \theta_1$  if  $\{L_i\}_{i=1}^K$  are identically distributed. For this, pick any measurable set  $\mathcal{A}$  and write

$$\begin{aligned} P[A_i \in \mathcal{A} | \theta = \theta_1] &= \int \int P[A_i \in \mathcal{A} | \theta = \theta_1, L_e = l_e, L_i = l_i] dP_{L_i}(l_i) dP_{L_e}(l_e) \\ &= \int \int v(\mathcal{A} | \theta_1 \xi(\mathbf{dist}(l_i, l_e))) dP_{L_i}(l_i) dP_{L_e}(l_e) \end{aligned} \quad (\text{C.11})$$

$$= \int \int v(\mathcal{A} | \theta_1 \xi(\mathbf{dist}(l_j, l_e))) dP_{L_j}(l_j) dP_{L_e}(l_e) = P[A_j \in \mathcal{A} | \theta = \theta_1], \quad (\text{C.12})$$

where (C.11) follows from (5.1) in Assumption V; and (C.12) follows because  $\{L_i\}_{i=1}^K$  are assumed i.i.d. ■

**Lemma C.2.** *Under the Assumptions II, III, IV, V, and VI of Chapter 5,  $\{A_i\}_{i=1}^K$  are conditionally independent given  $\{\theta = \theta_1, L_e = l_e\}$ ; i.e., for any  $l_e \in S_e$ ,*

$$P\left[\bigcap_{i=1}^K \{A_i \leq a_i\} \middle| \theta = \theta_1, L_e = l_e\right] = \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1, L_e = l_e]. \quad (\text{C.13})$$

If  $\{L_i\}_{i=1}^K$  are i.i.d. random variables, then  $\{A_i\}_{i=1}^K$  are conditionally i.i.d. given  $\{\theta = \theta_1, L_e = l_e\}$ .

*Proof:* Let  $P_{\mathbf{L}}$  be the probability measure of the random vector  $\mathbf{L} := (L_1, \dots, L_K)$  and write

$$\prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1, L_e = l_e] = \prod_{i=1}^K \int P[A_i \leq a_i | \theta = \theta_1, L_i = l_i, L_e = l_e] dP_{L_i}(l_i) \quad (\text{C.14})$$

$$= \int \prod_{i=1}^K P[A_i \leq a_i | \theta = \theta_1, L_i = l_i, L_e = l_e] dP_{\mathbf{L}}(\mathbf{l}) \quad (\text{C.15})$$

$$= \int P\left[\bigcap_{i=1}^K \{A_i \leq a_i\} \middle| \theta = \theta_1, \mathbf{L} = \mathbf{l}, L_e = l_e\right] dP_{\mathbf{L}}(\mathbf{l}) \quad (\text{C.16})$$

$$= P\left[\bigcap_{i=1}^K \{A_i \leq a_i\} \middle| \theta = \theta_1, L_e = l_e\right], \quad (\text{C.17})$$

where (C.15) follows because  $\{L_i\}_{i=1}^K$  are assumed independent and (C.16) follows from Assumption VI.

To reach the conclusion that  $\{A_i\}_{i=1}^K$  are conditionally i.i.d. given  $\{\theta = \theta_1, L_e = l_e\}$  if  $\{L_i\}_{i=1}^K$  are i.i.d., it remains to prove that, for all  $l_e$ ,  $\{A_i\}_{i=1}^K$  are identically distributed given  $\{\theta = \theta_1, L_e = l_e\}$  if  $\{L_i\}_{i=1}^K$  are identically distributed.



For this, write

$$\begin{aligned} P[A_i \leq a | \theta = \theta_1, L_e = l_e] &= \int P[A_i \leq a | \theta = \theta_1, L_e = l_e, L_i = l_i] dP_{L_i}(l_i) \\ &= \int v((-\infty, a] | \theta_1 \xi(\mathbf{dist}(l_e, l_i))) dP_{L_i}(l_i) \end{aligned} \quad (\text{C.18})$$

$$= \int v((-\infty, a] | \theta_1 \xi(\mathbf{dist}(l_e, l_j))) dP_{L_j}(l_j) = P[A_j \leq a | \theta = \theta_1, L_e = l_e], \quad (\text{C.19})$$

where (C.18) follows from (5.1) in Assumption V; and (C.19) follows because  $\{L_i\}_{i=1}^K$  are assumed i.i.d.  $\blacksquare$

**Lemma C.3.** *Adopt Assumptions II, III, IV, and V of Chapter 5. Let  $X_i := \xi(\mathbf{dist}(L_i, L_e))$ , and let  $P_{X_i|L_e=l_e}$  be the probability measure given by, for any  $x$ ,  $P[X_i \leq x | L_e = l_e] := P[\xi(\mathbf{dist}(L_i, l_e)) \leq x]$ , which exists and defines a probability measure for all  $l_e \in S_e$ . If*

$$\forall x, \forall l_{e_1}, l_{e_2} \in S_e^*, P[X_i \leq x | L_e = l_{e_1}] = P[X_i \leq x | L_e = l_{e_2}]; \quad (\text{C.20})$$

*i.e., the distribution of  $X_i$  conditioned on  $L_e = l_e$  does not change as  $l_e$  varies over  $S_e^*$ , then*

$$\forall l_{e_1}, l_{e_2} \in S_e^*, \forall a, P[A_i \leq a | \theta = \theta_1, L_e = l_{e_1}] = P[A_i \leq a | \theta = \theta_1, L_e = l_{e_2}]; \quad (\text{C.21})$$

*i.e.,  $P[A_i \leq a | \theta = \theta_1, L_e = l_e]$  does not change as  $l_e$  varies over  $S_e^*$  for all  $a$ .*

*Proof:* Use Assumption V to define the function  $v^*$  as

$$P[A_i \leq a_i | \theta = \theta_1, L_i = l_i, L_e = l_e] = v((-\infty, a_i] | \theta_1 \cdot \xi(\mathbf{dist}(l_i, l_e))) =: v^*(a_i, \xi(\mathbf{dist}(l_i, l_e))), \quad (\text{C.22})$$

and write

$$\begin{aligned} P[A_i \leq a_i | \theta = \theta_1, L_e = l_{e_1}] &= \int P[A_i \leq a_i | \theta = \theta_1, L_e = l_{e_1}, X_i = x] dP_{X_i|\theta=\theta_1, L_e=l_{e_1}}(x) \\ &= \int v^*(a_i, x) dP_{X_i|L_e=l_{e_1}}(x) \end{aligned} \quad (\text{C.23})$$

$$= \int_0^\infty P[X_i \in \{x : v^*(a_i, x) > t\} | L_e = l_{e_1}] dt \quad (\text{C.24})$$

$$= \int_0^\infty P[X_i \in \{x : v^*(a_i, x) > t\} | L_e = l_{e_2}] dt \quad (\text{C.25})$$

$$= P[A_i \leq a_i | \theta = \theta_1, L_e = l_{e_2}] \quad (\text{C.26})$$

where (C.23) follows from Assumption V and because  $X_i$  does not depend on  $\theta$ , and (C.25) follows because (C.20) implies that  $P_{X_i|L_e=l_{e_1}}[\mathcal{B}] = P_{X_i|L_e=l_{e_2}}[\mathcal{B}]$  for any measurable set  $\mathcal{B}$ .  $\blacksquare$

**Lemma C.4.** *Adopt Assumptions II, III, IV, and V of Chapter 5. Let  $D_i := \mathbf{dist}(L_i, L_e)$ , and let  $P_{D_i|L_e=l_e}$  be the probability measure given by, for any  $d$ ,  $P[D_i \leq d | L_e = l_e] := P[\mathbf{dist}(L_i, l_e) \leq d]$ , which exists and defines a probability measure for all  $l_e \in S_e$ . Let  $X_i := \xi(\mathbf{dist}(L_i, L_e))$ , and let  $P_{X_i|L_e=l_e}$  be the probability measure given by, for any  $x$ ,*

$P[X_i \leq x | L_e = l_e] := P[\xi(\mathbf{dist}(L_i, l_e)) \leq x]$ , which exists and defines a probability measure for all  $l_e \in S_e$ . For any given  $S_e^* \subseteq S_e$ , if

$$\forall d, \forall l_{e_1}, l_{e_2} \in S_e^*, P[D_i \leq d | L_e = l_{e_1}] = P[D_i \leq d | L_e = l_{e_2}]; \quad (\text{C.27})$$

i.e., the distribution of  $D_i$  conditioned on  $L_e = l_e$  does not change as  $l_e$  varies over  $S_e^*$ , then

$$\forall x, \forall l_{e_1}, l_{e_2} \in S_e^*, P[X_i \leq x | L_e = l_{e_1}] = P[X_i \leq x | L_e = l_{e_2}]; \quad (\text{C.28})$$

i.e., the distribution of  $X_i$  conditioned on  $L_e = l_e$  does not change as  $l_e$  varies over  $S_e^*$ .

*Proof:* Note that  $X_i$  is a function of  $D_i$ ; thus, for any measurable set  $\mathcal{B}$ ,  $P_{X_i}[\mathcal{B}] := P_{D_i}[X_i^{-1}[\mathcal{B}]]$ , where  $X_i^{-1}[\mathcal{B}] := \{d : \xi(d) \in \mathcal{B}\}$ . In particular,  $\xi(\mathbf{dist}(L_i, l_e))$  is a function of  $\mathbf{dist}(L_i, l_e)$ ; thus,  $P_{X_i|L_e=l_e}[\mathcal{B}] := P_{D_i|L_e=l_e}[X_i^{-1}[\mathcal{B}]]$ .

Since (C.27) means that  $P_{D_i|L_e=l_{e_1}}[\mathcal{A}] = P_{D_i|L_e=l_{e_2}}[\mathcal{A}]$  for any measurable set  $\mathcal{A}$ , it implies that,  $\forall x, \forall l_{e_1}, l_{e_2} \in S_e^*, P[X_i \leq x | L_e = l_{e_1}] = P[X_i \leq x | L_e = l_{e_2}]$ . ■

**Lemma C.5.** Let  $\mathcal{H}$  be the set of all orthogonal  $\delta$  – by –  $\delta$  matrices. Let  $S_e$  and  $S_s$  be two given sets. For any given  $S_e^* \subseteq S_e$ , choose any

$$\mathcal{H}^* \subseteq \{H \in \mathcal{H} | HS_e^* = S_e^*\}, \quad (\text{C.29})$$

where  $HS_e^* := \{Hl | l \in S_e^*\}$ . If

- for any  $l_{e_1}, l_{e_2} \in S_e^*$ , there exists  $H_{1,2} \in \mathcal{H}^*$  for which  $H_{1,2}l_{e_1} = l_{e_2}$ ;
- for any  $H \in \mathcal{H}^*$ ,  $HS_s = S_s$ ;
- For any  $i$ ,  $L_i$  is distributed on  $S_s$  and has a distribution invariant to any  $H \in \mathcal{H}^*$ ; i.e., for any measurable set  $\mathcal{B} \subset \mathbb{R}^\delta$ ,  $\forall H \in \mathcal{H}^*$ ,  $P[L_i \in \mathcal{B}] = P[L_i \in H\mathcal{B}]$ ;

then the distribution of  $D_i := \|L_i - L_e\|$  conditioned on  $L_e = l_e$  is invariant to changes in  $l_e \in S_e^*$ ; i.e.,

$$\forall i, \forall d, \forall l_{e_1}, l_{e_2} \in S_e^*, P[D_i \leq d | L_e = l_{e_1}] = P[D_i \leq d | L_e = l_{e_2}]. \quad (\text{C.30})$$

*Proof:* To show that the distribution of  $D_i$  conditioned on  $L_e = l_e$  is invariant to changes in  $l_e \in S_e^*$ , pick any two elements  $l_{e_1}, l_{e_2} \in S_e^*$ , observe that  $D_i$  conditioned on  $L_e = l_e$  is equal to  $\|L_i - l_e\|$ , and write

$$\forall d, P[\|L_i - l_{e_1}\| \leq d] = P[L_i \in B_d(l_{e_1}) \cap S_s] \quad (\text{C.31})$$

$$= P[L_i \in H_{1,2}(B_d(l_{e_1}) \cap S_s)] \quad (\text{C.32})$$

$$= P[L_i \in H_{1,2}B_d(l_{e_1}) \cap H_{1,2}S_s] \quad (\text{C.33})$$

$$= P[L_i \in B_d(H_{1,2}l_{e_1}) \cap H_{1,2}S_s] \quad (\text{C.34})$$

$$= P[L_i \in B_d(l_{e_2}) \cap S_s] \quad (\text{C.35})$$

$$= P[\|L_i - l_{e_2}\| \leq d] \quad (\text{C.36})$$

where (C.32) is reached because it is assumed that there exists  $H_{1,2} \in \mathcal{H}^*$  for which  $H_{1,2}l_{e_1} = l_{e_2}$  and that  $L_i$  has a distribution invariant to any  $H \in \mathcal{H}^*$ , (C.33) follows because any matrix  $H \in \mathcal{H}$  is invertible, which means that  $H$  forms a bijective transformation and, therefore,  $H(S_1 \cap S_2) = H(S_1) \cap H(S_2)$ , (C.34) follows because any matrix  $H \in \mathcal{H}$  preserves distances since  $\|\cdot\|$  is invariant to orthogonal transformations [51], and (C.35) follows from the assumptions on  $\mathcal{H}^*$ . ■

**Lemma C.6.** *Let  $S \subset \mathbb{R}^\delta$  and let  $\mathcal{H}$  be the set of all orthogonal  $\delta - \text{by} - \delta$  matrices. If  $S = B_r(0)$  for any  $r > 0$ , where  $B_r(0)$  represent the closed ball centered at the origin and with radius  $r$ , or  $S = \partial B_r(0)$ , where  $\partial B_r(0)$  represents the boundary of  $B_r(0)$ , then  $HS = S$  for all  $H \in \mathcal{H}$ .*

*Proof:* Assume first  $S = \partial B_r(0)$ . To show that  $HS = S$ , pick  $l \in HS \Leftrightarrow H^{-1}l \in S \Leftrightarrow \|H^{-1}l\| = r$ . Recalling that the norm operator  $\|l\| := \sqrt{l^T l}$  is invariant to orthogonal transformations [51] and that  $H^{-1}$  is itself an orthogonal transformation, it follows that  $\|H^{-1}l\| = r \Leftrightarrow \|l\| = r \Leftrightarrow l \in S$ . The case  $S = B_r(0)$  follows from the same rationale, replacing  $\|\cdot\| = r$  with  $\|\cdot\| \leq r$ . ■

**Lemma C.7.** *Let  $S_e \subset \mathbb{R}^2$  and assume*

- $S_e$  is a regular convex polygon with  $n$  vertices and a circumscribing circle  $\partial B_{r_M}(0)$  centered at the origin and with radius  $r_M$ ;
- $S_e^* = c_1(S_e \cap \partial B_{r_M}(0))$  for any  $c_1 \in [0, 1]$ ; i.e.,  $S_e^*$  is the set formed by all the vertices of the regular polygon multiplied by a scaling factor  $c_1$ .
- $S_s = B_r(0)$  for any  $r > 0$  or  $S_s = c_2 \cdot S_e$ , for any  $c_2 > 0$ ; i.e.,  $S_s$  is either a disk or  $S_s$  is a regular convex polygon with a circumscribing circle centered at the origin and with radius  $c_2 \cdot r_M$ ;

then the set of transformations  $\mathcal{H}^* := \{H_j : j \in \{0, 1, \dots, n-1\}\}$ , with  $H_j$  given by

$$H_j := \begin{bmatrix} \cos(2\pi j/n) & \sin(2\pi j/n) \\ -\sin(2\pi j/n) & \cos(2\pi j/n) \end{bmatrix}, \quad (\text{C.37})$$

satisfies

1. for any  $H_j \in \mathcal{H}^*$ ,  $H_j S_e^* = S_e^*$ ,  $H_j S_e = S_e$ , and  $H_j S_s = S_s$ ;
2. for any  $l_{e_1}, l_{e_2} \in S_e^*$ , there exists  $H_{1,2} \in \mathcal{H}^*$  for which  $H_{1,2}l_{e_1} = l_{e_2}$ .

*Proof:* A polygon with  $n$  vertices is defined as the set of line segments

$$\mathcal{V} := \{v_0 v_1, v_1 v_2, \dots, v_{n-1} v_0\} \quad (\text{C.38})$$

between vertices  $\{v_i\}_{i=0}^{n-1}$  and can therefore be represented by the ordered list of vertices  $\mathcal{V}_o := [v_0, \dots, v_{n-1}]$  [34] and note that the same regular polygon would result if one applies a circular shift operation in the ordered list of vertices since they result in the same set of line segments  $\mathcal{V}$ . Assume without loss of generality that the vertices of the regular polygon are located at  $v_i = (r_M \cos(2\pi i/n), r_M \sin(2\pi i/n))$  [34].

In order to show that  $\forall H_j \in \mathcal{H}^*$ ,  $H_j$  satisfies  $H_j S_e^* = S_e^*$  and  $H_j S_e = S_e$ , it is enough to show that  $H_j$  applied to the ordered list of vertices in  $\mathcal{V}_o$  corresponds to a circular shift in the ordered list of vertices, which defines the same set of line segments  $\mathcal{V}$  and, therefore, the same regular polygon that defines both  $S_e^*$  and  $S_e$ . To show that  $H_j \mathcal{V}_o$  results in a circular shift in  $\mathcal{V}_o$ , pick any  $v_i$  and observe that,  $H_j v_i = (r_M \cos(2\pi(i-j)/n), r_M \sin(2\pi(i-j)/n))$ . This means that, for all  $i \geq j$ ,  $H_j v_i = v_{i-j}$  and, for all  $i < j$ ,  $H_j v_i = (r_M \cos(2\pi(n+i-j)/n), r_M \sin(2\pi(n+i-j)/n)) = v_{n+i-j}$ . Thus,  $H_j \mathcal{V}_o$  results in the circularly shifted ordered list  $[v_{n-j}, v_{n+1-j}, \dots, v_0, v_1, \dots, v_{n-1-j}]$ . Note that  $H_j S_e^* = S_e^*$  for any value of  $c_1 \in [0, 1]$  since  $c_1(S_e \cap \partial B_{r_M}(0))$  are the vertices of a regular polygon with  $n$  vertices as well.

For the same reasons,  $H_j S_s = S_s$  if  $S_s = c_2 \cdot S_e$ . For the case  $S_s = B_r(0)$ , observe that every  $H_j$  is an orthogonal matrix; i.e.,  $H_j \in \mathcal{H}$  of Lemma C.6. Thus,  $\forall H \in \mathcal{H}, HS_s = S_s$ .

To show the second conclusion, pick any two vertices  $v_{i_1}$  and  $v_{i_2}$  of the regular polygon. If  $i_1 \geq i_2$ , let  $H_{1,2} = H_{j_1}$  for  $j_1 = i_1 - i_2$ . Since  $H_j v_i = v_{i-j}$  when  $i \geq j$ , and  $i_1 \geq j_1$ ,  $H_{j_1} v_{i_1} = v_{i_1-j_1} = v_{i_2}$ . If  $i_1 < i_2$ , let  $H_{1,2} = H_{j_2}$  for  $j_2 = n + i_1 - i_2$ . Since  $H_j v_i = v_{n+i-j}$  when  $i < j$ , and  $i_1 < j_2$ ,  $H_{j_2} v_{i_1} = v_{n+i_1-j_2} = v_{i_2}$ . ■

## Appendix D: Example that Illustrates the Need for an Augmented Definition for a Least Favorable Distribution

Consider that under a simple hypothesis  $H_0$ , a random variable  $X$  has a symmetric p.d.f. given by

$$f_0(x) = \begin{cases} 0.05, & 0 \leq |x| < 2, \\ 0.2, & 2 \leq |x| < 4, \\ 0, & 4 \leq |x|; \end{cases} \quad (\text{D.1})$$

and consider that under a composite hypothesis  $H_1$ , the p.d.f.  $f_1(x) \in \mathcal{F}_3$ , where

$$\mathcal{F}_3 := \left\{ f_{1,c}(x) \mid c \in \{1/3, 1/2, 2/3\} \right\} \quad (\text{D.2})$$

and  $f_{1,c}(x) := c \cdot f_0(x-2) + (1-c) \cdot f_0(x-4)$ .

Assuming that  $X$  is available for the decision maker, a least favorable distribution using the definition of [73] will be computed. For this, one finds the distribution in  $\mathcal{F}_3$  for which the optimum Neyman–Pearson detector provides the lowest probability of detection for a given probability of false alarm; i.e., one finds the value of  $c^-$  that satisfies:

$$\forall c' \in \{1/3, 1/2, 2/3\}, \beta(c', \phi_{c'}) \geq \beta(c^-, \phi_{c^-}), \quad (\text{D.3})$$

where  $\phi_c$  denotes the Neyman–Pearson detector

$$\phi_c(x) = \begin{cases} 1, & \mathcal{L}_c(x) > t_c, \\ \gamma_c, & \mathcal{L}_c(x) = t_c, \\ 0, & \mathcal{L}_c(x) < t_c, \end{cases} \quad (\text{D.4})$$

where  $\mathcal{L}_c(x) = f_{1,c}(x)/f_0(x)$  if  $f_0(x) > 0$ , and  $\mathcal{L}_c(x) = \infty$  if  $f_0(x) = 0$ .

For each possible  $c$ , the values of  $\gamma_c$  and  $t_c$  were chosen to achieve a probability of false alarm given by  $\alpha_{\max} = 0.05$ , and the resulting (maximum) probability of detection  $\beta(c, \phi_c)$  was computed. Table D.1 shows the results of these computations:

$c$	$\gamma_c$	$t_c$	$\beta(c, \phi_c)$	$\beta(c, \phi_{2/3})$
1/3	0.5	3	$37/60 \approx 0.6167$	$8/15 \approx 0.5333$
1/2	0.5	2.5	0.575	0.55
2/3	0.5	8/3	$17/30 \approx 0.5667$	$17/30 \approx 0.5667$

**Table D.1.** Randomization parameter and decision threshold for the Neyman–Pearson detectors to satisfy a probability of false alarm equal to 0.05, probability of detection of the Neyman–Pearson detectors, and probability of detection when using the detector  $\phi_{2/3}$  for each  $c$ .

As the designer examines the results in the column  $\beta(c, \phi_c)$  in the Table D.1, it concludes that the least favorable distribution using the definition of [73] is  $f_{1,2/3}(x)$ .

Table D.1 also shows in its last column the resulting probability of detection for each possible distribution  $f_{1,c}(x)$  if the designer chooses the detector  $\phi_{2/3}(x)$ . It is possible to see that the probability of detection decreases if the actual distribution  $f_{1,c}(x)$  is different from the least favorable distribution under the definition of [73].

This example illustrates that designs under the least favorable distribution according to the definition of [73] may not ensure a minimum probability of detection for a given probability of false alarm.

## Appendix E: Considerations About the Conditional Distribution of the Measurements

**Lemma E.1.** *Adopt the additive model of (2.1) and Assumptions V and VIII of Chapter 5. Let  $S'_e \subset S_e$  be either a subset that satisfies  $P[L_e \in S'_e] > 0$  or  $S'_e = \{l_e\}$  for any  $l_e \in S_e$ . Let  $Z_i$  assume values in some  $\mathcal{Z} \subset \mathbb{R}$  and let  $\mathcal{Z}_+ := \mathcal{Z} \cap \mathbb{R}_+$ . Let  $A_i$  assume values in  $\mathcal{Z}_+$  and  $W_i$  assume values in  $\mathcal{Z}$ . Then*

- *For any measurable set  $\mathcal{B}$ ,  $P[Z_i \in \mathcal{B} | \theta = \theta_1, L_e \in S'_e]$  is a probability measure;<sup>1</sup>*
- *The conditional distribution  $P_{Z_i | \theta=0}$  has a c.d.f.  $F_{Z_i | \theta=0}(z) = F_W(z)$  and a density function  $f_{Z_i | \theta=0}(z) = f_W(z)$  with respect to the measure  $\mu$ ;*
- *The conditional distribution  $P_{Z_i | \theta=\theta_1, L_e \in S'_e}$  has a c.d.f. and a density function with respect to the measure  $\mu$  respectively given by*

$$F_{Z_i | \theta=\theta_1, L_e \in S'_e}(z) = \int F_W(z-a) dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a) \quad (\text{E.1})$$

$$f_{Z_i | \theta=\theta_1, L_e \in S'_e}(z) = \int f_W(z-a) dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a). \quad (\text{E.2})$$

*Proof:* For either case of  $S'_e$ ,  $P_{A_i | \theta=\theta_1, L_e \in S'_e}$  is a probability measure. To see this, recall that, when  $S'_e$  is such that  $P[L_e \in S'_e] > 0$ , the conditioning on  $\{\theta = \theta_1, L_e \in S'_e\}$  becomes an elementary conditional probability, which always defines a probability measure [42]. For the case  $S'_e = \{l_e\}$ ,  $P_{A_i | \theta=\theta_1, L_e \in S'_e}$  is a probability measure for all  $\theta_1$  and all  $l_e$  because of Assumption V.

Let  $Y = \mathbf{1}\{L_e \in S'_e\}$ , observe that  $P_{A_i | \theta=\theta_1, L_e \in S'_e}$  and  $P_{A_i | \theta=\theta_1, Y=1}$  define the same probability measure for either case of  $S'_e$ , and write <sup>2</sup>

$$\begin{aligned} P[Z_i \in \mathcal{B} | \theta = \theta_1, L_e \in S'_e] &= P[Z_i \in \mathcal{B} | \theta = \theta_1, Y = 1] \\ &= \int P[Z_i \in \mathcal{B} | \theta = \theta_1, Y = 1, A_i = a] dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a). \end{aligned} \quad (\text{E.3})$$

From the additive model of (2.1),  $Z_i$  given  $A_i = a$  does not depend on  $L_e$  and  $P[Z_i \in \mathcal{B} | \theta = \theta_1, Y = 1, A_i = a] = P[Z_i \in \mathcal{B} | \theta = \theta_1, A_i = a]$  for either case of  $S'_e$ . Using this in (E.3), it is possible to write

$$P[Z_i \in \mathcal{B} | \theta = \theta_1, L_e \in S'_e] = \int P[Z_i \in \mathcal{B} | \theta = \theta_1, A_i = a] dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a) \quad (\text{E.4})$$

$$= \int \int P[Z_i \in \mathcal{B} | \theta = \theta_1, A_i = a, W_i = w] dP_W(w) dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a) \quad (\text{E.5})$$

$$= \int \int \mathbf{1}\{w+a \in \mathcal{B}\} dP_W(w) dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a) \quad (\text{E.6})$$

$$= \int_{\{w+a \in \mathcal{B}\}} dP_W(w) dP_{A_i | \theta=\theta_1, L_e \in S'_e}(a), \quad (\text{E.7})$$

<sup>1</sup>For  $S'_e = \{l_e\}$ , recall that  $P[Z_i \in \mathcal{B} | \theta = \theta_1, L_e = l_e]$  always exists as a measurable function of  $\theta_1$  and  $l_e$ ; however, it does not always form a probability measure [42].

<sup>2</sup>The introduction of the random variable  $Y$  is to make the integrand of (E.3) well defined.

and, observing that the set  $\{W + A_i \in \mathcal{B}\}$  is measurable in the smallest sigma-algebra that makes  $W$  and  $A_i$  measurable, (E.7) defines the probability measure of  $P[Z_i \in \mathcal{B} | \theta = \theta_1, L_e \in S'_e]$  for either case of  $S'_e$ .

The conditional distribution and density of  $Z_i$  given  $\theta = 0$  follow because  $Z_i = W_i$  when  $\theta = 0$ .

To show (E.1), just take  $\mathcal{B} = \mathcal{Z} \cap (-\infty, z]$ , recall that  $\theta_1 > 0$ , and (E.1) follows from (E.6).

To show (E.2) for either case of  $S'_e$ , use (E.1) with  $\mathcal{B} = \mathcal{Z} \cap (-\infty, z]$  and write

$$\begin{aligned} P[Z_i \leq z | \theta = \theta_1, L_e \in S'_e] &= \int_{\mathcal{Z}_+} \int_{\{x \in \mathcal{Z} : x \leq z - a\}} f_W(x) d\mu(x) dP_{A_i | \theta = \theta_1, L_e \in S'_e}(a) \\ &= \int_{\mathcal{Z}_+} \int_0^\infty \mu(\{x \in \mathcal{Z} : \mathbf{1}\{x \leq z - a\} f_W(x) > t\}) dt dP_{A_i | \theta = \theta_1, L_e \in S'_e} \\ &= \int_{\mathcal{Z}_+} \int_0^\infty \mu(\{y \in \mathcal{Z} : \mathbf{1}\{y \leq z\} f_W(y - a) > t\}) dt dP_{A_i | \theta = \theta_1, L_e \in S'_e} \end{aligned} \quad (\text{E.8})$$

$$= \int_{\mathcal{Z}_+} \int_{\{y \in \mathcal{Z} : y \leq z\}} f_W(y - a) d\mu(y) dP_{A_i | \theta = \theta_1, L_e \in S'_e} \quad (\text{E.9})$$

$$= \int_{\{y \in \mathcal{Z} : y \leq z\}} \int_{\mathcal{Z}_+} f_W(y - a) dP_{A_i | \theta = \theta_1, L_e \in S'_e} d\mu(y), \quad (\text{E.10})$$

where (E.8) follows because, for any set  $\mathcal{B}$ ,  $\mu(\mathcal{B}) = \mu(\mathcal{B} + a)$  for any  $a$  when  $\mu$  is Lebesgue measure and for any  $a \in \mathbb{Z}$  when  $\mu$  is counting measure on  $\mathbb{Z}$ , and (E.10) follows from the Fubini–Tonelli theorem [47]. ■



## Appendix F: Supporting Propositions for Chapter 7

**Lemma F.1.** Consider the conditions and the distributed detection system specified in Proposition 7.6. Let  $U^{\max} = \max\{U_{i,y}^{\max} : i \in \{1, \dots, K\}, y \in \{1, \dots, M\}\}$ .

- For any  $i$ , any  $m \in \{1, \dots, M\}$ , and any  $u \in \{0, \dots, U^{\max}\}$ ,

$$P[U_i = u | \theta = 0, Y = m] = 0 \Rightarrow P[U_i = u | \theta = \theta_1, Y = m] = 0. \quad (\text{F.1})$$

- If the thresholds that define the sensor decision intervals were such that  $\exists u^- \in \{0, \dots, U_{i,m}^{\max}\}$  with  $P[U_i = u^- | \theta = 0, Y = m] = 0$  for some  $i$  and some  $m$ ; then it is possible to change the sensor function so that it uses one less sensor output and  $P[U_i = u | \theta = 0, Y = m] > 0$  for all  $u \in \{0, \dots, U_{i,m}^{\max} - 1\}$  without changing the probability of detection or the probability of false alarm at the fusion center.

*Proof:* To reach (F.1), recall that it is assumed that  $P_{Z_i|\theta=\theta_1}$  is absolutely continuous with respect to  $P_{Z_i|\theta=0}$ ; thus, since  $P[U_i = u | \theta = 0, Y = m] = 0$  implies  $P_{Z_i|\theta=0}\{z : \mathcal{L}_{Z_i, S'_e}(z) \in I_{i,m,u}\} = 0$ , it follows that  $0 = P_{Z_i|\theta=\theta_1}\{z : \mathcal{L}_{Z_i, S'_e}(z) \in I_{i,m,u}\} = P[U_i = u | \theta = \theta_1, Y = m]$ .

To prove the second result, assume that there exists a sensor  $i^-$  for which the sensor function  $\phi_{i^-}$  follows the form of (7.33) but condition (7.35) is not satisfied for a single  $m^-$  and a single  $u^-$ ; i.e., the intervals  $\{I_{i^-,m^-,u}\}_{u=0}^{U_{i^-,m^-}^{\max}}$  are such that there exists a value  $u^-$  such that  $P[U_{i^-} = u^- | \theta = 0, Y = m^-] = 0$ . The strategy to prove the second result is to show that it is possible to build a new set of intervals with one less interval; i.e.,  $\{I'_{i^-,m^-,u}\}_{u=0}^{U_{i^-,m^-}^{\max}-1}$ , in which the interval  $I_{i^-,m^-,u^-}$  is either excluded or made part of another interval without changing the system performance.

Consider the alternative sensor function  $\phi'_{i^-}(z_{i^-}, y)$  that satisfies:  $\forall m \neq m^-, \phi'_{i^-}(z_{i^-}, m) = \phi_{i^-}(z_{i^-}, m)$ , and for  $m = m^-$ ,

$$\phi'_{i^-}(z_{i^-}, m^-) = \sum_{u=0}^{U_{i^-,m^-}^{\max}-1} u \cdot \mathbf{1}\left\{\mathcal{L}_{Z_{i^-}, S'_e}(z_{i^-}) \in I'_{i^-,m^-,u}\right\} \quad (\text{F.2})$$

where  $\forall u < u^-$ ,  $I'_{i^-,m^-,u} = I_{i^-,m^-,u}$ ; if  $u^- \leq U_{i^-,m^-}^{\max} - 1$ , then  $I'_{i^-,m^-,u^-} = I_{i^-,m^-,u^-} \cup I_{i^-,m^-,u^-+1}$ ; and if  $u^- \leq U_{i^-,m^-}^{\max} - 2$ , then  $\forall u \in \{u^- + 1, \dots, U_{i^-,m^-}^{\max} - 1\}$ ,  $I'_{i^-,m^-,u} = I_{i^-,m^-,u+1}$ .

To see that the alternative sensor function  $\phi'_{i^-}(z_{i^-}, y)$  satisfies condition (7.35), consider first the case  $u^- = U_{i^-,m^-}^{\max}$ ; in this case, condition (7.35) is satisfied because the sensor  $\phi'_{i^-}(z_{i^-}, m^-)$  has  $U_{i^-,m^-}^{\max} - 1$  as maximum output and it is assumed that  $P[U'_{i^-} = u | \theta = 0, Y = m^-] > 0$  for all  $u < u^-$ . For the case  $u^- < U_{i^-,m^-}^{\max}$ , let  $U'_{i^-} = \phi'_{i^-}(z_{i^-}, y)$  and note that  $P[U'_{i^-} = u^- | \theta = 0, Y = m^-] > 0$  because  $I'_{i^-,m^-,u^-} \supset I_{i^-,m^-,u^-+1}$  and  $P[\mathcal{L}_{Z_{i^-}, S'_e}(z_{i^-}) \in I_{i^-,m^-,u^-+1} | \theta = 0] > 0$  by assumption on  $\phi_{i^-}$ . Since  $\forall u \in \{u^- + 1, \dots, U_{i^-,m^-}^{\max} - 1\}$ ,  $I'_{i^-,m^-,u} = I_{i^-,m^-,u+1}$  and the only value of  $u$  in which condition (7.35) was not respected is  $u^-$ , the alternative sensor function  $\phi'_{i^-}(z_{i^-}, y)$  satisfies condition (7.35).

The goal is to show that  $P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_j] = P[\phi'_0(\mathbf{U}, Y) = 1 | \theta = \theta_j]$  for either  $\theta_j = 0$  or  $\theta_j = \theta_1$ , where  $\phi'_0$  is the fusion function given by (7.36) with  $\mathcal{L}_{U'_{i^-}, S'_{e,m^-}}(u_i)$  given by (7.37).

Since  $P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_j] = \sum_{m=1}^M P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_j, Y = m] P[Y = m]$  and the two sensor functions are identical for  $m \neq m^-$ , it is enough to show that  $P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_j, Y = m^-] = P[\phi'_0(\mathbf{U}, Y) = 1 | \theta = \theta_j, Y = m^-]$ .

To show this, write

$$\begin{aligned} P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_j, Y = m^-] &= \sum_{u=0}^{U_{i^-, m^-}^{\max}} P[\phi_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U_{i^-} = u] P[U_{i^-} = u | \theta = \theta_j, Y = m^-] \\ &= \sum_{u=0}^{u^- - 1} P[\phi_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U_{i^-} = u] P[U_{i^-} = u | \theta = \theta_j, Y = m^-] \end{aligned} \quad (\text{F.3})$$

$$+ \sum_{u=u^-+1}^{U_{i^-, m^-}^{\max}} P[\phi_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U_{i^-} = u] P[U_{i^-} = u | \theta = \theta_j, Y = m^-] \quad (\text{F.4})$$

where the term corresponding to  $u^-$  is not needed in (F.4) when  $\theta_j = 0$  because  $P[U_{i^-} = u^- | \theta = 0, Y = m^-] = 0$ . To see that it is also not needed for  $\theta_j = 1$ , recall that (F.1) gives that  $P[U_{i^-} = u^- | \theta = \theta_1, Y = m^-] = 0$  as well.

Similarly, write

$$\begin{aligned} P[\phi'_0(\mathbf{U}, Y) = 1 | \theta = \theta_j, Y = m^-] &= \sum_{u=0}^{U_{i^-, m^-}^{\max} - 1} P[\phi'_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U'_{i^-} = u] P[U'_{i^-} = u | \theta = \theta_j, Y = m^-] \\ &= \sum_{u=0}^{u^- - 1} P[\phi'_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U'_{i^-} = u] P[U'_{i^-} = u | \theta = \theta_j, Y = m^-] \end{aligned} \quad (\text{F.5})$$

$$+ \sum_{u=u^-}^{U_{i^-, m^-}^{\max} - 1} P[\phi'_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U'_{i^-} = u] P[U'_{i^-} = u | \theta = \theta_j, Y = m^-]. \quad (\text{F.6})$$

In order to show that the term in (F.5) equals the term in (F.3), and the term in (F.6) equals the term in (F.4), rewrite  $\phi_0(\mathbf{U}, m^-)$  as  $\phi_0(\mathbf{U}, m^-) = \mathbf{1}\{g(\mathbf{U}^{(i^-)}, m^-) \mathcal{L}_{U_{i^-}, S'_{e, m^-}}(u_{i^-}) \in I_{0, m^-}\}$ , where  $\mathbf{U}^{(i^-)}$  is the vector of  $\mathbf{U}$  without  $U_{i^-}$  and  $g(\mathbf{U}^{(i^-)}, m^-) := \prod_{i \neq i^-} \mathcal{L}_{U_i, S'_{e, m^-}}(U_i)$ . With this notation,

$$\begin{aligned} P[\phi_0(\mathbf{U}, Y) = 1 | \theta = \theta_j, Y = m^-, U_{i^-} = u] \\ = P[\mathbf{1}\{g(\mathbf{U}^{(i^-)}, m^-) \mathcal{L}_{U_{i^-}, S'_{e, m^-}}(u) \in I_{0, m^-}\} = 1 | \theta = \theta_j, Y = m^-], \end{aligned} \quad (\text{F.7})$$

and note that  $\mathcal{L}_{U_{i^-}, S'_{e, m^-}}(u)$  is no longer random.

To show that the term in (F.5) equals the term in (F.3), it is shown that  $\forall u < u^-$ ,

$$\begin{aligned} P[\phi_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U_{i^-} = u] P[U_{i^-} = u | \theta = \theta_j, Y = m^-] \\ = P[\phi'_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U'_{i^-} = u] P[U'_{i^-} = u | \theta = \theta_j, Y = m^-], \end{aligned} \quad (\text{F.8})$$

which will be shown by proving that  $\forall u < u^-$ ,  $\mathcal{L}_{U_{i^-}, S'_{e, m^-}}(u) = \mathcal{L}'_{U'_{i^-}, S'_{e, m^-}}(u)$  and  $P[U_{i^-} = u | \theta = \theta_j, Y = m^-] = P[U'_{i^-} = u | \theta = \theta_j, Y = m^-]$ . Note that, since  $\mathcal{L}_{U_{i^-}, S'_{e, m^-}}(u)$  is just a constant multiplying  $g(\mathbf{U}^{(i^-)}, m^-)$ , if  $\mathcal{L}_{U_{i^-}, S'_{e, m^-}}(u)$  equals  $\mathcal{L}'_{U'_{i^-}, S'_{e, m^-}}(u)$ , then both  $\phi_0(\mathbf{U}, m^-)$  and  $\phi'_0(\mathbf{U}, m^-)$  will have the same distribution.

To show that  $P[U_{i^-} = u | \theta = \theta_j, Y = m^-] = P[U'_{i^-} = u | \theta = \theta_j, Y = m^-]$ , write

$$\begin{aligned} P[U_{i^-} = u | \theta = \theta_j, Y = m^-] &= P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I_{i^-, m^-, u} | \theta = \theta_j] \\ &= P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I'_{i^-, m^-, u} | \theta = \theta_j] = P[U'_{i^-} = u | \theta = \theta_j, Y = m^-], \end{aligned} \quad (\text{F.9})$$

which follows because  $I'_{i^-, m^-, u} = I_{i^-, m^-, u}$  for all  $u < u^-$ .

To show that  $\mathcal{L}_{U_{i^-}, S'_e, m^-}(u) = \mathcal{L}_{U'_{i^-}, S'_e, m^-}(u)$ , recall that it is assumed that both  $P[U_{i^-} = u | \theta = 0, Y = m^-] > 0$  and  $P[U'_{i^-} = u | \theta = 0, Y = m^-] > 0$  for  $u < u^-$ , which means that, from (7.37),

$$\begin{aligned} \mathcal{L}_{U_{i^-}, S'_e, m^-}(u) &= \frac{P[U_{i^-} = u | \theta = \theta_1, L_e \in S'_e, Y = m^-]}{P[U_{i^-} = u | \theta = 0, Y = m^-]} \\ &= \frac{P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I_{i^-, m^-, u} | \theta = \theta_1, L_e \in S'_e]}{P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I_{i^-, m^-, u} | \theta = \theta_0]} \\ &= \frac{P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I'_{i^-, m^-, u} | \theta = \theta_1, L_e \in S'_e]}{P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I'_{i^-, m^-, u} | \theta = \theta_0]} \end{aligned} \quad (\text{F.10})$$

$$\begin{aligned} &= \frac{P[U'_{i^-} = u | \theta = \theta_1, L_e \in S'_e, Y = m^-]}{P[U'_{i^-} = u | \theta = 0, Y = m^-]} \\ &= \mathcal{L}_{U'_{i^-}, S'_e, m^-}(u), \end{aligned} \quad (\text{F.11})$$

where (F.10) follows because  $I'_{i^-, m^-, u} = I_{i^-, m^-, u}$  for all  $u < u^-$ . With (F.9) and (F.11), it is proven that the term in (F.5) equals the term in (F.3).

To show that the term in (F.6) equals the term in (F.4), consider first the case  $u^- = U_{i^-, m^-}^{\max}$ . In this case, both term in (F.6) and the term in (F.4) equal 0 since there are no terms under the summation. For the case  $u^- < U_{i^-, m^-}^{\max}$ , it will be shown that  $\forall u \in \{u^-, \dots, U_{i^-, m^-}^{\max} - 1\}$ ,

$$\begin{aligned} P[\phi_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U_{i^-} = u + 1] &P[U_{i^-} = u + 1 | \theta = \theta_j, Y = m^-] = \\ P[\phi'_0(\mathbf{U}, m^-) = 1 | \theta = \theta_j, Y = m^-, U'_{i^-} = u] &P[U'_{i^-} = u | \theta = \theta_j, Y = m^-], \end{aligned} \quad (\text{F.12})$$

which, as before, will be shown by proving that  $\mathcal{L}_{U_{i^-}, S'_e, m^-}(u + 1) = \mathcal{L}_{U'_{i^-}, S'_e, m^-}(u)$  and  $P[U_{i^-} = u + 1 | \theta = \theta_j, Y = m^-] = P[U'_{i^-} = u | \theta = \theta_j, Y = m^-]$  for all the specified  $u$ .

To show that  $\forall u \in \{u^-, \dots, U_{i^-, m^-}^{\max} - 1\}$ ,  $P[U_{i^-} = u + 1 | \theta = \theta_j, Y = m^-] = P[U'_{i^-} = u | \theta = \theta_j, Y = m^-]$ , write

$$\begin{aligned} P[U_{i^-} = u + 1 | \theta = \theta_j, Y = m^-] &= P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I_{i^-, m^-, u+1} | \theta = \theta_j] \\ &= P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I'_{i^-, m^-, u} | \theta = \theta_j] = P[U'_{i^-} = u | \theta = \theta_j, Y = m^-], \end{aligned} \quad (\text{F.13})$$

which follows because  $I'_{i^-, m^-, u} = I_{i^-, m^-, u+1}$  for all  $u \in \{u^- + 1, \dots, U_{i^-, m^-}^{\max} - 1\}$ . To see that (F.13) also follows for  $u = u^-$ , recall that  $I'_{i^-, m^-, u^-} = I_{i^-, m^-, u^-} \cup I_{i^-, m^-, u^-+1}$  and

$$\begin{aligned} P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I'_{i^-, m^-, u^-} | \theta = \theta_j] &= P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I_{i^-, m^-, u^-} | \theta = \theta_j] \\ &\quad + P[\mathcal{L}_{Z_{i^-}, S'_e}(Z_{i^-}) \in I_{i^-, m^-, u^-+1} | \theta = \theta_j]. \end{aligned} \quad (\text{F.14})$$

The first term in the sum of (F.14) equals 0 because, by assumption,  $P[U_{i-} = u^- | \theta = 0, Y = m^-] = 0$  and (F.1) gives that  $P[U_{i-} = u^- | \theta = \theta_j, Y = m^-] = 0$  as well. Therefore,  $P[U_{i-}' = u^- | \theta = \theta_j, Y = m^-] = P[U_{i-} = u^- + 1 | \theta = \theta_j, Y = m^-]$ .

To show that  $\forall u \in \{u^-, \dots, U_{i-, m^-}^{\max} - 1\}$ ,  $\mathcal{L}_{U_{i-}, S_e', m^-}(u+1) = \mathcal{L}_{U_{i-}, S_e', m^-}(u)$ , observe that it is only required to consider the case  $\mathcal{L}_{U_{i-}, S_e', m^-}(u+1) < \infty$ , because, by assumption,  $P[U_{i-} = u | \theta = 0, Y = m^-] = 0$  only for  $u^-$ , and write

$$\begin{aligned} \mathcal{L}_{U_{i-}, S_e', m^-}(u+1) &= \frac{P[U_{i-} = u+1 | \theta = \theta_1, L_e \in S_e', Y = m^-]}{P[U_{i-} = u+1 | \theta = 0, Y = m^-]} \\ &= \frac{P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u+1} | \theta = \theta_1, L_e \in S_e']}{P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u+1} | \theta = 0]} \end{aligned} \quad (\text{F.15})$$

$$= \frac{P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u} | \theta = \theta_1, L_e \in S_e']}{P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u} | \theta = 0]} \quad (\text{F.16})$$

$$\begin{aligned} &= \frac{P[U_{i-}' = u | \theta = \theta_1, L_e \in S_e', Y = m^-]}{P[U_{i-}' = u | \theta = 0, Y = m^-]} \\ &= \mathcal{L}_{U_{i-}, S_e', m^-}(u), \end{aligned} \quad (\text{F.17})$$

where (F.16) follows because  $I_{i-, m^-, u} = I_{i-, m^-, u+1}$  for all  $u > u^-$ . To see that (F.16) also follows for  $u = u^-$ , recall that (F.1) indicates that the assumption  $P[U_{i-} = u^- | \theta = 0, Y = m^-] = 0$  implies  $P[U_{i-} = u^- | \theta = \theta_1, Y = m^-] = 0$ . From the definition of  $U_{i-}$ , this implies  $P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u^-} | \theta = \theta_1, L_e \in S_e'] = P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u^-} | \theta = 0] = 0$ . Thus,  $P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u^-} | \theta = \theta_1, L_e \in S_e']$  can be added to the numerator of (F.15) and  $P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u^-} | \theta = 0]$  can be added to the denominator of (F.15). The term in (F.15) for  $u = u^-$  then becomes equal to

$$\frac{P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u^-+1} \cup I_{i-, m^-, u^-} | \theta = \theta_1, L_e \in S_e']}{P[\mathcal{L}_{Z_{i-}, S_e'}(Z_{i-}) \in I_{i-, m^-, u^-+1} \cup I_{i-, m^-, u^-} | \theta = 0]} \quad (\text{F.18})$$

and (F.16) follows because  $I_{i-, m^-, u^-} = I_{i-, m^-, u^-} \cup I_{i-, m^-, u^-+1}$ .

With (F.13) and (F.17), it is proven that the term in (F.6) equals the term in (F.4) and the second conclusion of this proposition is reached.  $\blacksquare$

**Lemma F.2.** *Under the conditions of in Proposition 7.7, if*

$$\forall t, P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e, Y = m] \geq P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e^-, Y = m], \quad (\text{F.19})$$

then

$$\forall t, P\left[\prod_{i=1}^K T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e, Y = m\right] \geq P\left[\prod_{i=1}^K T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e^-, Y = m\right]. \quad (\text{F.20})$$

*Proof:* This Lemma is essentially Lemma 2 of [123]. It is included here in the notation of Proposition 7.7 to facilitate the understanding and for completeness.

Because of Assumption V,  $P[T_{0,i}(U_i, m) > t | \theta = \theta_1, L_e = l_e, Y = m]$  forms a probability measure for all  $l_e$  and all  $m$ . Thus, for all  $i$ , one can build two random variables  $R_i^*$  and  $S_i^*$  such that

$$\forall x, \quad P[R_i^* \leq x] := P[T_{0,i}(U_i, m) \leq x | \theta = \theta_1, L_e = l_e^-, Y = m] \quad (\text{F.21})$$

$$P[S_i^* \leq x] := P[T_{0,i}(U_i, m) \leq x | \theta = \theta_1, L_e = l_e, Y = m] \quad (\text{F.22})$$

and it is possible to impose  $R_i^* \geq 0$  and  $S_i^* \geq 0$  because  $T_{0,i}(U_i, m) \geq 0$  as well.

Let  $R_i := \log R_i^*$  and  $S_i := \log S_i^*$ ; let  $F_{0,i}(x) := P[R_i \leq x]$  and  $F_{1,i}(x) := P[S_i \leq x]$ ; and note that Condition (F.19) indicates that  $\forall x, F_{1,i}(x) \leq F_{0,i}(x)$ .

Lemma 3.4.1 of [73] gives that, if  $\forall x, F_{1,i}(x) \leq F_{0,i}(x)$ , then there exists one random variable  $V_i$  and 2 real functions  $f_{0,i}(v)$  and  $f_{1,i}(v)$  such that  $\forall v, f_{0,i}(v) \leq f_{1,i}(v)$ ; and  $f_{0,i}(V_i)$  is a random variable with  $F_{0,i}(v)$  as c.d.f. and  $f_{1,i}(V_i)$  is a random variable with  $F_{1,i}(v)$  as c.d.f.

Since the above is true for all  $i$ , for any vector  $(v_1, \dots, v_K)$ ,  $\sum_{i=1}^K f_{0,i}(v_i) \leq \sum_{i=1}^K f_{1,i}(v_i)$ ; and  $\sum_{i=1}^K f_{0,i}(V_i) \leq \sum_{i=1}^K f_{1,i}(V_i)$  pointwise; i.e., this inequality is valid for any realization of  $\{V_i\}_{i=1}^K$ ; and it follows that

$$\forall x, P \left[ \sum_{i=1}^K f_{0,i}(V_i) \leq x \right] \geq P \left[ \sum_{i=1}^K f_{1,i}(V_i) \leq x \right]. \quad (\text{F.23})$$

Since  $f_{0,i}(V_i)$  and  $R_i$  have the same distribution for all  $i$ ,  $\sum_{i=1}^K f_{0,i}(V_i)$  and  $\sum_{i=1}^K R_i$  have the same distribution. Similarly,  $\sum_{i=1}^K f_{1,i}(V_i)$  and  $\sum_{i=1}^K S_i$  have the same distribution, and (F.23) is equivalent to

$$\begin{aligned} \forall x, P \left[ \sum_{i=1}^K R_i \leq x \right] &\geq P \left[ \sum_{i=1}^K S_i \leq x \right] \Leftrightarrow \forall x, P \left[ \sum_{i=1}^K \log R_i^* \leq x \right] \geq P \left[ \sum_{i=1}^K \log S_i^* \leq x \right] \\ &\Leftrightarrow \forall x', P \left[ \prod_{i=1}^K R_i^* \leq x' \right] \geq P \left[ \prod_{i=1}^K S_i^* \leq x' \right]. \end{aligned} \quad (\text{F.24})$$

Since the distribution of  $R_i^*$  is equal to the distribution of  $T_{0,i}(U_i, m)$  conditioned on  $\{\theta = \theta_1, L_e = l_e^-, Y = m\}$  for every  $i$ , it follows that  $\forall x', P[\prod_{i=1}^K R_i^* \leq x'] = P[\prod_{i=1}^K T_{0,i}(U_i, m) \leq x' | \theta = \theta_1, L_e = l_e^-, Y = m]$ ; and using the same reasoning for the product of  $\{S_i^*\}_{i=1}^K$ , (F.24) implies (F.20). ■

**Lemma F.3.** For any two random variables  $V_1$  and  $V_2$ :

$$\forall v, P[V_1 \geq v] \geq P[V_2 \geq v] \Leftrightarrow \forall v, P[V_1 > v] \geq P[V_2 > v].$$

*Proof:* ( $\Rightarrow$ ):  $\forall v$ , let  $v_n$  be a decreasing sequence with  $v_n \rightarrow v$  and  $v_n > v, \forall n$ . Thus, since measures are continuous from below [47],  $P_{V_j}\{[v_n, \infty)\} \rightarrow P_{V_j}\{\cup_{n=1}^\infty [v_n, \infty)\}$ , as  $n \rightarrow \infty$  for  $j \in \{1, 2\}$ . Since  $\cup_{n=1}^\infty [v_n, \infty) = (v, \infty)$ ,  $\lim_{n \rightarrow \infty} P_{V_1}\{[v_n, \infty)\} - P_{V_2}\{[v_n, \infty)\} = P_{V_1}\{(v, \infty)\} - P_{V_2}\{(v, \infty)\}$ . From the left-hand side,  $P_{V_1}\{[v_n, \infty)\} - P_{V_2}\{[v_n, \infty)\} \geq 0$  for all  $n$ ; thus, the limit is also nonnegative and the right-hand side follows. ( $\Leftarrow$ ):  $\forall v$ , let  $v_n$  be an increasing sequence with  $v_n \rightarrow v$  and  $v_n < v, \forall n$ . From the right-hand side,  $P_{V_1}\{(-\infty, v_n]\} - P_{V_2}\{(-\infty, v_n]\} \leq 0$  and proceed as before to reach  $\lim_{n \rightarrow \infty} P_{V_1}\{(-\infty, v_n]\} - P_{V_2}\{(-\infty, v_n]\} = P_{V_1}\{(-\infty, v)\} - P_{V_2}\{(-\infty, v)\} \leq 0$ . ■

**Lemma F.4.** Let  $S_e \subset \mathbb{R}^2$  and  $S_s = c \cdot S_e$  or  $S_s = c \cdot B_{r_M}(0)$ , for any  $c > 0$ . Assume that either  $S_e = B_{r_M}(0)$  or  $S_e$  is a regular convex polygon with  $n$  vertices and a circumscribing circle  $\partial B_{r_M}(0)$  for some  $r_M > 0$ . Let  $B_r^\circ(l) := B_r(l) - \partial B_r(l)$ ; i.e.,  $B_r^\circ(l)$  is the open ball of radius  $r$  and centered at  $l$ . If  $l_e^- \in \partial B_{r_M}(0) \cap S_e$  and  $l_e^+ = 0$ , then, for any  $l_e \in S_e$  and any  $r \geq 0$ ,

$$m(B_r^\circ(l_e^-) \cap S_s) \leq m(B_r^\circ(l_e) \cap S_s) \leq m(B_r^\circ(l_e^+) \cap S_s), \quad (\text{F.25})$$

where  $m(\mathcal{B})$  represents Lebesgue measure of  $\mathcal{B}$  on  $\mathbb{R}^2$ .

*Proof:* Observe that  $m(B_r^\circ(l_e) \cap S_s) = m(B_r(l_e) \cap S_s)$  and it suffices to prove (F.25) with closed balls instead of open balls.

Consider first the case  $S_e = B_{r_M}(0)$ . Since  $S_e$ ,  $S_s$ , and Lebesgue measure are invariant to orthogonal transformations, as shown during the proof of Proposition 6.2, it follows that, for any  $l_{e_1} \in S_e$  and any orthogonal matrix  $H$ , the point  $l_{e_2} := Hl_{e_1}$  is also an element of  $S_e$  and

$$m(B_r(l_{e_1}) \cap S_s) = m(HB_r(l_{e_1}) \cap HS_s) = m(B_r(Hl_{e_1}) \cap S_s) = m(B_r(l_{e_2}) \cap S_s). \quad (\text{F.26})$$

For any  $r > 0$ , pick any  $l_{e_1} \in S_e$ , and let  $l_{e_2} = H_\pi l_{e_1}$ , where  $H_\pi$  is the matrix that rotates  $l_{e_1}$  by  $\pi$ ; i.e., using the definition for  $H_j$  in (C.37) with  $n = 2$ ,  $H_\pi = H_1$ .

Since (F.26) is valid for  $H_\pi$  as well, applying Lemma F.5 for the chosen  $l_{e_1}$  and  $l_{e_2}$ , and observing that  $l_e^+ = 0$  is contained in the line segment that connects  $l_{e_1}$  and  $l_{e_2}$ , conclude that

$$m(B_r(l_{e_1}) \cap S_s) \leq m(B_r(l_e^+) \cap S_s), \quad (\text{F.27})$$

and since (F.27) is valid for all  $r > 0$  and all  $l_{e_1} \in S_e$ , the upper bound of (F.25) is proven.

To prove the lower bound of (F.25), pick any  $r > 0$  and any  $l_{e_3} \in S_e$  and observe that (F.26) is valid when  $l_{e_1}, l_{e_2} \in \partial B_{r_M}(0)$  are such that the line segment  $\ell$  that connects  $l_{e_1}$  and  $l_{e_2}$  contains both the origin and  $l_{e_3}$ . Thus, apply Lemma F.5 once again to conclude that,

$$m(B_r(l_{e_1}) \cap S_s) \leq m(B_r(l_{e_3}) \cap S_s), \quad (\text{F.28})$$

and since the same argument is valid for any  $l_{e_3} \in S_e$ , the lower bound of (F.25) is proven for any  $l_e^- \in \partial B_{r_M}(0)$ .

Consider now the case in which  $S_e$  is a regular convex polygon with  $n$  vertices and a circumscribing circle  $\partial B_{r_M}(0)$ .

In order to prove the upper bound of (F.25), pick any  $r > 0$  and define

$$S_M := \{l \in S_e : m(B_r(l) \cap S_s) \geq m(B_r(l_e) \cap S_s), \forall l_e \in S_e\}; \quad (\text{F.29})$$

and the goal is to prove that  $l_e^+ = 0 \in S_M$ .

Observe that  $S_M$  is nonempty because  $S_e$  is closed and bounded, and therefore compact, and because the function  $g(l) := m(B_r(l) \cap S_s)$  is continuous on  $S_e$ , there exists a point  $l^+ \in S_e$  such that  $g(l^+) = \sup_{l \in S_e} \{g(l)\}$ . To see why

$g$  is continuous at any  $l_e \in S_e$ , pick a sequence  $l_n \rightarrow l_e$  and define  $h_n(l) := \mathbf{1}\{l \in B_r(l_n) \cap S_s\}$ . Observe that  $h_n(l) \rightarrow h(l) := \mathbf{1}\{l \in B_r(l_e) \cap S_s\}$  for all  $l \notin \partial(B_r(l_e) \cap S_s)$ ; i.e., for all  $l$  not on the boundary of  $B_r(l_e) \cap S_s$ . This means that  $h_n(l) \rightarrow h(l)$  almost-everywhere in Lebesgue measure in  $\mathbb{R}^2$  since  $m(\partial(B_r(l_e) \cap S_s)) = 0$ . Observe also that  $h_n(l) \leq \mathbf{1}\{l \in S_s\} \in L_1(m)$  because  $m(S_s) < \infty$ . Thus, one can use the Lebesgue Dominated Convergence Theorem to conclude that  $g(l_n) := \int h_n(l) dm(l) \rightarrow \int h(l) dm(l) =: g(l_e)$ .

Since  $S_M$  is nonempty, pick an  $l_e \in S_M$ . If such  $l_e = l_e^+$ , the conclusion is reached. If  $l_e \neq 0$ , then it is claimed here that there exists  $n$  points  $l_0 = l_e, l_1, \dots, l_{n-1}$  within  $S_e$ , such that  $m(B_r(l_e) \cap S_s) = m(B_r(l_j) \cap S_s)$  for  $j = 1, \dots, n-1$ , and because  $l_e \in S_M, l_j \in S_M$  for  $j = 1, \dots, n-1$ . To see this, note that  $S_e$  and  $S_s$  satisfy the conditions of Lemma C.7, define  $l_j := H_j l_e$  with  $H_j$  given by (C.37), and write

$$m(B_r(l_e) \cap S_s) = m(H_j(B_r(l_e) \cap S_s)) = m(H_j B_r(l_e) \cap H_j S_s) \quad (\text{F.30})$$

$$= m(B_r(H_j l_e) \cap H_j S_s) = m(B_r(l_j) \cap S_s) \quad (\text{F.31})$$

where the first equality of (F.30) follows because Lebesgue measure is invariant to orthogonal transformations, the second equality (F.30) follows because any orthogonal matrix is invertible, which means that  $H_j$  forms a bijective transformation and, therefore, for any two sets  $S_1$  and  $S_2$ ,  $H_j(S_1 \cap S_2) = H_j(S_1) \cap H_j(S_2)$ ; the first equality of (F.31) follows because any orthogonal matrix preserves distances, and the second equality of (F.31) follows from the definition of  $l_j$  and Lemma C.7.

From the set of  $n$  points  $l_0 = l_e, l_1, \dots, l_{n-1}$ , form the set of line segments  $\mathcal{V}_1 := \{l_0 l_1, l_1 l_2, \dots, l_{n-1} l_0\}$ .

Since  $m(B_r(l_i) \cap S_s) = m(B_r(l_j) \cap S_s)$  for any line  $l_i l_j \in \mathcal{V}_1$ , Lemma F.5 can be applied to obtain that  $m(B_r(l) \cap S_s) \geq m(B_r(l_j) \cap S_s)$  for any  $l$  within the line  $l_i l_j$ ; and, since both  $l_i$  and  $l_j$  are elements of  $S_M$ , obtain that  $m(B_r(l) \cap S_s) = m(B_r(l_j) \cap S_s)$  and  $l \in S_M$  for any  $l$  within the line  $l_i l_j$ . Since  $l_i$  is also part of another line segment in  $\mathcal{V}_1$ ,  $m(B_r(l) \cap S_s) = m(B_r(l_i) \cap S_s)$  for any  $l$  in such a line segment as well. Repeating this rationale, it follows that  $m(B_r(l) \cap S_s) = m(B_r(l_e) \cap S_s)$  for any point  $l \in \{l^* \in l_i l_j : l_i l_j \in \mathcal{V}_1\}$  and all of such points belong to  $S_M$ .

The convex hull of the set  $\{l_0 = l_e, l_1, \dots, l_{n-1}\}$  contains  $l_e^+ = 0$  because these points were obtained through the rotation operations  $H_j$ .

Pick any  $l'$  and  $l''$  that belong to  $\{l^* \in l_i l_j : l_i l_j \in \mathcal{V}_1\}$  such that the line segment  $l' l''$  contains  $l_e^+ = 0$ . Since  $m(B_r(l') \cap S_s) = m(B_r(l'') \cap S_s)$ , Lemma F.5 gives that  $m(B_r(l_e^+) \cap S_s) \geq m(B_r(l') \cap S_s)$ . Since  $\mathcal{V}_1 \subset S_M, l' \in S_M$  and  $l'' \in S_M$ , and it follows that  $l_e^+ \in S_M$  as well.

In order to prove the lower bound of (F.25), pick any  $l_e \in S_e$  and obtain the line segment that contains  $l_e$  and connects  $l_e^+ = 0$  and the point  $l'$  on the boundary of  $S_e$ . Since  $m(B_r(l_e^+) \cap S_s) \geq m(B_r(l') \cap S_s)$ , the second result of Lemma F.5 gives that

$$m(B_r(l_e) \cap S_s) \geq m(B_r(l') \cap S_s). \quad (\text{F.32})$$

Observe now that  $l'$  belongs to a line that connects two vertices of the regular polygon. Let  $l_i^*$  and  $l_j^*$  denote these vertices and note that both  $l_i^*$  and  $l_j^*$  belong to  $\partial B_{r_M}(0) \cap S_e$ . Since  $m(B_r(l_i^*) \cap S_s) = m(B_r(l_j^*) \cap S_s)$ , the first result of

Lemma F.5 yields

$$m(B_r(l') \cap S_s) \geq m(B_r(l_i^*) \cap S_s) = m(B_r(l_e^-) \cap S_s), \quad (\text{F.33})$$

and the lower bound of (F.25) follows from (F.32) and (F.33).  $\blacksquare$

**Lemma F.5.** *Let  $S_s \subset \mathbb{R}^2$  be a closed, bounded, nonempty, convex set. For any measurable set  $\mathcal{B} \subset \mathbb{R}^2$ , let  $m(\mathcal{B})$  represents Lebesgue measure of  $\mathcal{B}$  on  $\mathbb{R}^2$ . Pick any  $r \geq 0$  and any  $l_{e_1}, l_{e_2} \in S_s$ . If either*

$$m(B_r(l_{e_1}) \cap S_s) = m(B_r(l_{e_2}) \cap S_s) \quad (\text{F.34})$$

or

$$\forall l_e \in S_s, m(B_r(l_{e_1}) \cap S_s) \geq m(B_r(l_e) \cap S_s), \quad (\text{F.35})$$

then

$$\forall \lambda \in [0, 1], m(B_r((1-\lambda)l_{e_1} + \lambda l_{e_2}) \cap S_s) \geq m(B_r(l_{e_2}) \cap S_s). \quad (\text{F.36})$$

*Proof:* In order to prove (F.36), the Brunn–Minkowski theorem [49] shall be used.<sup>1</sup> For that, define

$$M_\lambda := B_r((1-\lambda)l_{e_1} + \lambda l_{e_2}) \cap S_s \quad (\text{F.37})$$

$$M_\lambda^* := \{l^* : l^* = (1-\lambda)l' + \lambda l'', \text{ where } l' \in B_r(l_{e_1}) \cap S_s, l'' \in B_r(l_{e_2}) \cap S_s\}; \quad (\text{F.38})$$

observe that

$$\begin{aligned} M_0 &= B_r(l_{e_1}) \cap S_s \quad \text{and} \quad M_1 = B_r(l_{e_2}) \cap S_s. \\ M_0^* &= M_0 \quad \text{and} \quad M_1^* = M_1; \end{aligned} \quad (\text{F.39})$$

and observe that

$$M_\lambda^* = (1-\lambda)M_0^* \oplus \lambda M_1^*, \quad (\text{F.40})$$

where the operator  $\oplus$  denotes the vector sum; i.e., for any two sets  $S_1$  and  $S_2$  in  $\mathbb{R}^\delta$ ,  $S_1 \oplus S_2 := \{s_1 + s_2 : s_1 \in S_1, s_2 \in S_2\}$ .

It is now claimed that  $M_\lambda^* \subseteq M_\lambda$ . To see this, pick any  $l^* \in M_\lambda^*$  and it is first shown that  $l^* \in B_r((1-\lambda)l_{e_1} + \lambda l_{e_2})$ . Since  $l^* \in M_\lambda^*$ , there are  $l' \in B_r(l_{e_1}) \cap S_s$ , and  $l'' \in B_r(l_{e_2}) \cap S_s$ , such that  $l^* = (1-\lambda)l' + \lambda l''$ . Since  $l' \in B_r(l_{e_1}) \cap S_s$ , the vector  $\Delta_1 := l' - l_{e_1}$  must satisfy  $\|\Delta_1\| \leq r$ ; and, similarly,  $\Delta_2 := l'' - l_{e_2}$  must satisfy  $\|\Delta_2\| \leq r$ . Thus, it follows that

$$\begin{aligned} l^* &= (1-\lambda)l' + \lambda l'' = (1-\lambda)(l_{e_1} + \Delta_1) + \lambda(l_{e_2} + \Delta_2) \\ &= (1-\lambda)l_{e_1} + \lambda l_{e_2} + (1-\lambda)\Delta_1 + \lambda\Delta_2, \end{aligned} \quad (\text{F.41})$$

---

<sup>1</sup>This proof was inspired by [48].



and since  $\|(1-\lambda)\Delta_1 + \lambda\Delta_2\| \leq r$ , it follows that  $\|l^* - (1-\lambda)l_{e_1} - \lambda l_{e_2}\| \leq r$ ; thus,  $l^* \in B_r((1-\lambda)l_{e_1} + \lambda l_{e_2})$ . To see that  $l^* \in S_s$ , observe that both  $l', l'' \in S_s$  and  $S_s$  is convex; thus, since  $l^*$  belongs to the line segment formed by  $l'$  and  $l''$ ,  $l^* \in S_s$  as well. It then follows that  $l^* \in M_\lambda$ .

Further observe that both  $M_0^*$  and  $M_1^*$  are closed, bounded, nonempty convex sets. To see this, note that both  $B_r(l_e)$  and  $S_s$  are convex sets and recall that the intersection of convex sets results in a convex set [19]. Since  $B_r(l_e)$  and  $S_s$  are closed, bounded, and nonempty sets, and since  $l_e \in S_s$ , then  $M_0^*$  and  $M_1^*$  are also closed, bounded, and nonempty sets.

Since  $M_0^*$  and  $M_1^*$  are closed, bounded, nonempty convex sets, one can apply the Brunn–Minkowski theorem [49, p. 361], which gives that

$$\sqrt{m((1-\lambda)M_0^* \oplus \lambda M_1^*)} \geq (1-\lambda)\sqrt{m(M_0^*)} + \lambda\sqrt{m(M_1^*)} \quad (\text{F.42})$$

$$= (1-\lambda)\sqrt{m(M_0)} + \lambda\sqrt{m(M_1)} \quad (\text{F.43})$$

$$= (1-\lambda)\sqrt{m(B_r(l_{e_1}) \cap S_s)} + \lambda\sqrt{m(B_r(l_{e_2}) \cap S_s)} \quad (\text{F.44})$$

where  $m$  denotes Lebesgue measure in  $\mathbb{R}^2$ , (F.43) and (F.44) follow from (F.39); and, using (F.40) in (F.44), one obtains

$$\sqrt{m(M_\lambda^*)} \geq (1-\lambda)\sqrt{m(B_r(l_{e_1}) \cap S_s)} + \lambda\sqrt{m(B_r(l_{e_2}) \cap S_s)}. \quad (\text{F.45})$$

If  $m(B_r(l_{e_1}) \cap S_s) = m(B_r(l_{e_2}) \cap S_s)$ ; i.e., if condition (F.34) is satisfied, then (F.45) becomes

$$\sqrt{m(M_\lambda^*)} \geq \sqrt{m(B_r(l_{e_2}) \cap S_s)}, \quad (\text{F.46})$$

and it follows that  $m(M_\lambda^*) \geq m(B_r(l_{e_2}) \cap S_s)$ . Thus, since  $M_\lambda^* \subseteq M_\lambda$ , it follows that  $m(M_\lambda) \geq m(M_\lambda^*) \geq m(B_r(l_{e_2}) \cap S_s)$ , and the conclusion (F.36) follows from (F.37).

Consider now the case:  $\forall l_e \in S_s, m(B_r(l_{e_1}) \cap S_s) \geq m(B_r(l_e) \cap S_s)$ ; i.e., condition (F.35) is satisfied. Let  $l' := (1-\lambda)l_{e_1} + \lambda l_{e_2}$  and note that  $l' \in S_s$  since  $S_s$  is convex. Thus, condition (F.35) means that

$$\begin{aligned} m(B_r(l_{e_1}) \cap S_s) &\geq m(B_r(l') \cap S_s) \\ &= m(B_r((1-\lambda)l_{e_1} + \lambda l_{e_2}) \cap S_s) = m(M_\lambda). \end{aligned} \quad (\text{F.47})$$

Using (F.47) in (F.45), one obtains

$$\begin{aligned} \sqrt{m(M_\lambda^*)} &\geq (1-\lambda)\sqrt{m(B_r(l_{e_1}) \cap S_s)} + \lambda\sqrt{m(B_r(l_{e_2}) \cap S_s)} \\ &\geq (1-\lambda)\sqrt{m(M_\lambda)} + \lambda\sqrt{m(B_r(l_{e_2}) \cap S_s)} \end{aligned} \quad (\text{F.48})$$

$$\geq (1-\lambda)\sqrt{m(M_\lambda^*)} + \lambda\sqrt{m(B_r(l_{e_2}) \cap S_s)} \quad (\text{F.49})$$

where (F.49) follows because  $M_\lambda^* \subseteq M_\lambda$ . Rearranging the terms in (F.49), the conclusion (F.36) follows.  $\blacksquare$

**Proposition F.6.** *If  $A^-$  is a Poisson random variable with parameter  $x^- > 0$  and  $A^+$  is a Poisson random variable with parameter  $x^+ > x^-$ , then, for all  $a$ ,*

$$P[A^- > a] \leq P[A^+ > a] \quad (\text{F.50})$$

*Proof:* For  $a < 0$ , conclusion (F.50) is reached since both sides equal 1. For  $a \geq 0$ , let  $a^* := \lfloor a \rfloor$  and observe that, for any Poisson random variable  $A$  with parameter  $x$ ,  $P[A > a] = 1 - \sum_{n=0}^{a^*} x^n \exp\{-x\}/n!$ , which is a differentiable function of  $x$  and its derivative equals

$$\begin{aligned} \frac{d}{dx} \left( 1 - \sum_{n=0}^{a^*} \frac{x^n \exp\{-x\}}{n!} \right) &= \exp\{-x\} - \sum_{n=1}^{a^*} \left( \frac{nx^{n-1} \exp\{-x\}}{n!} - \frac{x^n \exp\{-x\}}{n!} \right) \\ &= \exp\{-x\} - \sum_{m=0}^{a^*-1} \frac{x^m \exp\{-x\}}{m!} + \sum_{n=1}^{a^*} \frac{x^n \exp\{-x\}}{n!} \\ &= - \sum_{m=1}^{a^*-1} \frac{x^m \exp\{-x\}}{m!} + \sum_{n=1}^{a^*} \frac{x^n \exp\{-x\}}{n!} = \frac{x^{a^*} \exp\{-x\}}{a^*!} > 0, \end{aligned}$$

and the conclusion (F.50) follows. ■

## Appendix G: Supporting Example

In this appendix, the performance of a simple sensor detection system is evaluated in order to illustrate some of the concepts discussed in this dissertation.

Assume the additive model (2.1) of Chapter 2 with  $A_i = \theta_1 \xi(\|L_i - L_e\|)$  with  $\xi(r) = \mathbf{1}\{r < r_0\}$  for some  $r_0 > 0$  and  $\theta_1 = 1$ . Let  $F_W$  be the distribution of the noise random variable  $W_i$ , which is assumed independent of  $\theta$ ,  $L_i$ , and  $L_e$ . Assume that a single sensor ( $K = 1$ ) is deployed in the sensor deployment region  $S_s = cS_e$  for some  $c > 0$ . Assume that the sensor location  $L_1$  is independent of  $\theta$  and  $L_e$  and is uniformly distributed in  $S_s$ . Let this simple sensor detection system be centralized with fusion function given by  $\phi_0(z_1, y_0) = \mathbf{1}\{z_1 > t_0\}$ . Let  $t_0$  be such that  $P[Z_1 > t_0 | \theta = 0] \leq \alpha_{\max}$  for some  $\alpha_{\max} \in (0, 1)$ .

Consider that the emitter location  $L_e$  has a distribution  $P_{L_e}^{(l_e)}$  that satisfies  $P[L_e = l_e] = 1$  for some  $l_e \in S_e$ .

Under the distribution  $P_{L_e}^{(l_e)}$ , the probability of detection  $\beta(P_{L_e}^{(l_e)}, \phi_0)$  is given by

$$\begin{aligned} \beta(P_{L_e}^{(l_e)}, \phi_0) &= P[Z_1 > t_0 | \theta = 1, L_1 \in (S_s \cap B_{r_0}^\circ(l_e))] \cdot P[L_1 \in (S_s \cap B_{r_0}^\circ(l_e)) | \theta = 1] + \\ &\quad P[Z_1 > t_0 | \theta = 1, L_1 \in (S_s \cap B_{r_0}^\circ(l_e)^c)] \cdot P[L_1 \in (S_s \cap B_{r_0}^\circ(l_e)^c) | \theta = 1], \end{aligned} \quad (\text{G.1})$$

where  $B_{r_0}^\circ(l_e)$  represents the open ball of radius  $r_0$  centered at  $l_e$ .

Because  $\xi(r) = \mathbf{1}\{r < r_0\}$ ,  $A_1 = 1$  only when  $L_1 \in (S_s \cap B_{r_0}^\circ(l_e))$ , which means that  $Z_1$  conditioned on the event  $\{\theta = 1, L_1 \in (S_s \cap B_{r_0}^\circ(l_e))\}$  satisfies  $Z_1 = 1 + W_1$ , and  $Z_1$  conditioned on the event  $\{\theta = 1, L_1 \in (S_s \cap B_{r_0}^\circ(l_e)^c)\}$  satisfies  $Z_1 = W_1$ . Since  $W_1$ ,  $\theta$ ,  $L_1$ , and  $L_e$  are independent,

$$\begin{aligned} \beta(P_{L_e}^{(l_e)}, \phi_0) &= P[1 + W_1 > t_0] \cdot P[L_1 \in (S_s \cap B_{r_0}^\circ(l_e))] + P[W_1 > t_0] \cdot (1 - P[L_1 \in (S_s \cap B_{r_0}^\circ(l_e))]) \\ &= P[L_1 \in (S_s \cap B_{r_0}^\circ(l_e))] \cdot [F_W(t_0) - F_W(t_0 - 1)] + (1 - F_W(t_0)) \\ &= \frac{m(S_s \cap B_{r_0}^\circ(l_e))}{m(S_s)} \cdot [F_W(t_0) - F_W(t_0 - 1)] + (1 - F_W(t_0)), \end{aligned} \quad (\text{G.2})$$

where (G.2) follows from the uniform distribution of  $L_1$ .

## Appendix H: Detailed Calculations for the Example in Section 8.1

In order to facilitate the reading of this appendix, the conditions of the example in Section 8.1 are repeated.

Adopt the additive model of (2.1) and Assumptions I, II, III, IV, V, VI, and VII of Chapter 5. Let  $\mathbf{dist}(L_i, L_e) = \|L_i - L_e\|$  in Assumption V. Let the region of interest  $S_e$  be a disk centered at the origin with radius  $r_M$ ; i.e.,  $S_e = B_{r_M}(0)$ . Let  $S_s = cS_e$  for some  $c \geq 1$ , and let  $\{L_i\}_{i=1}^K$  be i.i.d. with the uniform distribution in  $S_s$ . For any  $l_i$  and  $l_e$ , let  $A_i$  conditioned on  $L_i = l_i$  and  $L_e = l_e$  be Poisson-distributed with parameter  $\theta_1 \xi(\|l_i - l_e\|)$  with  $\xi(r) = A_{\max}/(1+r)^2$  and  $\theta_1 = 1$ . Let  $W_i$  be Poisson-distributed with parameter  $\lambda_W > 0$ . For each  $K$ , let  $\underline{\phi}^K$  be the binary distributed detection system with given binary sensor functions

$$\phi_{i,K}(z_i, y) = \mathbf{1}\{z_i > t_s\}, \quad (\text{H.1})$$

where  $t_s = \lambda_W$  for this example. Assume that  $K$  dedicated and error-free communication channels are available; and let the fusion function be

$$\phi_{0,K}(\mathbf{u}, y_0) = \mathbf{1}\left\{\sum_{i=1}^K u_i \in I_{0,y_0}\right\}, \quad (\text{H.2})$$

where  $I_{0,0} = (k, \infty]$ ,  $I_{0,1} = [k, \infty]$ , and  $y_0$  is the realization of the randomization random variable  $Y_0 \in \{0, 1\}$ . From Theorem 3.1, (H.2) defines the optimal fusion function for the set of sensor functions (H.1) when  $U_i$  are conditionally i.i.d. given  $\theta$ .

To compute the probability of false alarm, observe that, under  $\theta = 0$ , the measurements  $\{Z_i\}_{i=1}^K$ , and therefore  $\{U_i\}_{i=1}^K$ , are i.i.d. under any distribution for  $L_e$ ; the decision statistic is a Binomial random variable; and the probability of false alarm is given by

$$P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 0] = P[Y_0 = 0] \sum_{j=k+1}^K \binom{K}{j} p_0^j (1-p_0)^{K-j} + P[Y_0 = 1] \sum_{j=k}^K \binom{K}{j} p_0^j (1-p_0)^{K-j}, \quad (\text{H.3})$$

where

$$p_0 := P[U_i = 1 | \theta = 0] = P[W_i > t_s] = 1 - \sum_{n=0}^{\lfloor t_s \rfloor} \frac{\lambda_W^n \exp\{-\lambda_W\}}{n!}. \quad (\text{H.4})$$

Using (H.3) and (H.4), it is possible to design a system that has a probability of false alarm equal to  $\alpha_{\max}$  as follows: for each  $K$ , find  $k$  that satisfies

$$\sum_{j=k+1}^K \binom{K}{j} p_0^j (1-p_0)^{K-j} \leq \alpha_{\max} \leq \sum_{j=k}^K \binom{K}{j} p_0^j (1-p_0)^{K-j}; \quad (\text{H.5})$$

and use the resulting  $k$  to define the parameter of the Bernoulli random variable  $Y_0$  as

$$P[Y_0 = 1] = \frac{\alpha_{\max} - \sum_{j=k+1}^K \binom{K}{j} p_0^j (1-p_0)^{K-j}}{\binom{K}{k} p_0^k (1-p_0)^{K-k}}. \quad (\text{H.6})$$

To compute the probability of detection, write

$$P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 1] = \int_{S_e} P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 1, L_e = l_e] dP_{L_e}(l_e); \quad (\text{H.7})$$

and, because  $\{L_i\}_{i=1}^K$  are i.i.d.,  $\{U_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta = 1$  and  $L_e = l_e$ , as shown in Lemma C.2, one finds that

$$\begin{aligned} P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 1, L_e = l_e] &= P[Y_0 = 0] \cdot \sum_{j=k+1}^K \binom{K}{j} p_{1,l_e}^j (1 - p_{1,l_e})^{K-j} + \\ &P[Y_0 = 1] \cdot \sum_{j=k}^K \binom{K}{j} p_{1,l_e}^j (1 - p_{1,l_e})^{K-j}, \end{aligned} \quad (\text{H.8})$$

where

$$p_{1,l_e} := P[U_i = 1 | \theta = 1, L_e = l_e] = P[A_i + W_i > t_s | \theta = 1, L_e = l_e] = \frac{1}{m(S_s)} \int_0^{2\pi} \int_0^{c r_M} p_{1,l_e,l_i}(r, \omega) r dr d\omega; \quad (\text{H.9})$$

$$p_{1,l_e,l_i}(r, \omega) := 1 - \sum_{n=0}^{\lfloor t_s \rfloor} \frac{(\xi(\|l_e - (r \cos \omega, r \sin \omega)\|) + \lambda_W)^n \exp\{-(\xi(\|l_e - (r \cos \omega, r \sin \omega)\|) + \lambda_W)\}}{n!}; \quad (\text{H.10})$$

$m(S_s)$  represents the area of  $S_s$ ; and polar coordinates were used in (H.9), which follows from the uniform distribution of  $L_i$ .

From Proposition 7.7, a least favorable emitter location distribution for  $\underline{\phi}^K$  is the distribution  $P_{L_e}^-$  that satisfies  $P[L_e \in S_e^*] = 1$  with  $S_e^* = \partial B_{r_M}(0)$ ; i.e.,  $P_{L_e}^-$  places the emitter in the boundary of the disk with probability 1.

To compute the probability of detection under  $P_{L_e}^-$ , observe that, because the distribution  $P_{L_e}^-$  satisfies the conditions of Proposition 6.2,  $\{U_i\}_{i=1}^K$  are conditionally i.i.d. given  $\theta = 1$  and are invariant to changes in  $l_e$ ; thus, the probability of detection is given by (H.8) with  $l_e$  equal to any of the points in  $\partial B_{r_M}(0)$ .

From Proposition 7.7, a most favorable emitter location distribution for  $\underline{\phi}^K$  is the distribution  $P_{L_e}^+$  that satisfies  $P[L_e = (0, 0)] = 1$ ; and the probability of detection is given by (H.8) with  $l_e = (0, 0)$ .

Under either  $P_{L_e}^-$  or  $P_{L_e}^+$ , it is not necessary to compute the integral in (H.7) because the integrand is constant over the support of either  $P_{L_e}^-$  or  $P_{L_e}^+$ ; however, this is not the case when  $L_e$  is distributed according to other distributions, such as the uniform distribution  $P_{L_e}^\mu$  on  $S_e$ .

To compute the probability of detection under  $P_{L_e}^\mu$ , use polar coordinates to rewrite (H.7) as

$$\begin{aligned} P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 1] &= \\ \frac{1}{m(S_e)} \int_0^{2\pi} \int_0^{r_M} P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 1, L_e = (r_e \cos \omega_e, r_e \sin \omega_e)] r_e dr_e d\omega_e, \end{aligned} \quad (\text{H.11})$$

where  $m(S_e)$  is the area of  $S_e$ ; and  $P[\phi_{0,K}(\mathbf{U}, Y_0) = 1 | \theta = 1, L_e = (r_e \cos \omega_e, r_e \sin \omega_e)]$  is given by (H.8) with  $p_{1,l_e}$  given by (H.9) with  $l_e = (r_e \cos \omega_e, r_e \sin \omega_e)$ .

The result of the computations is shown in Section 8.1.

## Appendix I: Supporting Material for Chapter 8

### I.1 Comparing Required Number of Sensors for Prescribed Performance

Assume that  $P_{L_e}^-$  and  $P_{L_e}^+$  are a least and a most favorable distributions for each detection system of the sequence  $\Phi$  and let  $P_{L_e}'$  be any other distribution for  $L_e$ .

Assume that  $\Phi$  satisfies  $K^{(\Phi, P_{L_e}^-)} < \infty$ . In this section, it will be explained why

$$\forall P_{L_e}', 1 \leq \frac{K^{(\Phi, P_{L_e}^-)}}{K^{(\Phi, P_{L_e}')}} \leq \frac{K^{(\Phi, P_{L_e}^-)}}{K^{(\Phi, P_{L_e}^+)}}. \quad (\text{I.1})$$

From the definition of  $K^{(\Phi, P_{L_e}^-)}$ ,  $\alpha(\underline{\phi}^{K^{(\Phi, P_{L_e}^-)}}) \leq \alpha_{\max}$ , and since  $P_{L_e}^-$  is a least favorable distribution for  $\underline{\phi}^K$  for every  $K$ , one can write

$$\beta_{\min} \leq \beta(P_{L_e}^-, \underline{\phi}^{K^{(\Phi, P_{L_e}^-)}}) \leq \beta(P_{L_e}', \underline{\phi}^{K^{(\Phi, P_{L_e}^-)}}); \quad (\text{I.2})$$

which implies that  $K^{(\Phi, P_{L_e}^-)} \in \{K : \beta(P_{L_e}', \underline{\phi}^K) \geq \beta_{\min}, \alpha(\underline{\phi}^K) \leq \alpha_{\max}\}$  and, from (8.2),  $K^{(\Phi, P_{L_e}')} \leq K^{(\Phi, P_{L_e}^-)}$ . The same rationale is used to show that  $K^{(\Phi, P_{L_e}^+)} \leq K^{(\Phi, P_{L_e}')} \leq K^{(\Phi, P_{L_e}^-)}$ .

Consider now that  $\mathcal{C}$  is any class that satisfies Condition 7.1, assume that  $P_{L_e}^-$  is a least favorable distribution for  $\underline{\phi}_{P_{L_e}^-|\mathcal{C}}^K$  for every  $K$ ; and assume that  $\underline{\phi}_{P_{L_e}^+|\mathcal{C}}^K$  and  $P_{L_e}^+$  satisfy the conditions of Lemma 7.2 for every  $K$ ; i.e.,  $\underline{\phi}_{P_{L_e}^-|\mathcal{C}}^K$  is maximin for  $\mathcal{C}$  and  $\underline{\phi}_{P_{L_e}^+|\mathcal{C}}^K$  is maximax for  $\mathcal{C}$ . Further assume that  $K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)} < \infty$ . It is now shown that

$$\forall P_{L_e}', 1 \leq \frac{K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)}}{K^{(\Phi_{P_{L_e}'|\mathcal{C}}, P_{L_e}^-)}} \leq \frac{K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)}}{K^{(\Phi_{P_{L_e}^+|\mathcal{C}}, P_{L_e}^+)}}. \quad (\text{I.3})$$

First, observe that  $\alpha(\underline{\phi}_{P_{L_e}^-|\mathcal{C}}^K) \leq \alpha_{\max}$ ,  $\alpha(\underline{\phi}_{P_{L_e}'|\mathcal{C}}^K) \leq \alpha_{\max}$ , and  $\alpha(\underline{\phi}_{P_{L_e}^+|\mathcal{C}}^K) \leq \alpha_{\max}$  for any  $K$ , which follow from (7.1).

Regarding the probabilities of detection, observe that

$$\forall K, \forall P_{L_e}', \beta(P_{L_e}^-, \underline{\phi}_{P_{L_e}^-|\mathcal{C}}^K) \leq \beta(P_{L_e}', \underline{\phi}_{P_{L_e}^-|\mathcal{C}}^K) \leq \beta(P_{L_e}', \underline{\phi}_{P_{L_e}'|\mathcal{C}}^K) \leq \beta(P_{L_e}^+, \underline{\phi}_{P_{L_e}^+|\mathcal{C}}^K), \quad (\text{I.4})$$

where the first inequality follows because  $P_{L_e}^-$  is a least favorable distribution for  $\underline{\phi}_{P_{L_e}^-|\mathcal{C}}^K$ ; the second inequality follows because  $\underline{\phi}_{P_{L_e}'|\mathcal{C}}^K$  is optimum under  $P_{L_e}'$ , and the last inequality follows because  $\underline{\phi}_{P_{L_e}^+|\mathcal{C}}^K$  and  $P_{L_e}^+$  satisfy the conditions of Lemma 7.2. Thus, since (I.4) is valid for every  $K$ , the second inequality of (I.4) can be used with  $K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)}$  and it is possible to reach that

$$\beta_{\min} \leq \beta(P_{L_e}^-, \underline{\phi}_{P_{L_e}^-|\mathcal{C}}^{K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)}}) \leq \beta(P_{L_e}', \underline{\phi}_{P_{L_e}'|\mathcal{C}}^{K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)}}). \quad (\text{I.5})$$

Since (7.1) defines that  $\alpha(\underline{\phi}_{P_{L_e}'|\mathcal{C}}^K) \leq \alpha_{\max}$  for any  $K$ , it follows that

$$K^{(\Phi_{P_{L_e}^-|\mathcal{C}}, P_{L_e}^-)} \in \{K : \beta(P_{L_e}', \underline{\phi}_{P_{L_e}'|\mathcal{C}}^K) \geq \beta_{\min}, \alpha(\underline{\phi}_{P_{L_e}'|\mathcal{C}}^K) \leq \alpha_{\max}\} \quad (\text{I.6})$$

and, from (8.2),  $K^{(\Phi_{P'_{Le}|\mathcal{E}}, P'_{Le})} \leq K^{(\Phi_{P_{Le}^-|\mathcal{E}}, P_{Le}^-)}$ . The same rationale is used to show that  $K^{(\Phi_{P_{Le}^+|\mathcal{E}}, P_{Le}^+)} \leq K^{(\Phi_{P'_{Le}|\mathcal{E}}, P'_{Le})} \leq K^{(\Phi_{P_{Le}^-|\mathcal{E}}, P_{Le}^-)}$ .

## I.2 Details Regarding the ARE Metric

To see why the ARE metric can be given by the squared ratio of the efficacies; i.e., by (8.9), let  $\beta_{\min}$  be the prescribed minimum asymptotic probability of detection for a maximum asymptotic probability of false alarm given by  $\alpha_{\max}$ .

Consider a first sequence of detection systems  $\Phi_{(1)}$  with a certain efficacy  $\zeta(\Phi_{(1)})$ . The asymptotic probability of detection of such a sequence of detectors depends on both the value  $c$  and on the efficacy  $\zeta(\Phi_{(1)})$ , as defined in Theorem 8.1. In order to have the asymptotic probability of detection (8.8) be equal to  $\beta_{\min}$ , it is necessary to adjust the constant  $c$ . Let  $c^{(1)}$  be such a constant.

Similarly, for a second sequence of detection systems  $\Phi_{(2)}$ , let  $c^{(2)}$  be the constant required to have its asymptotic probability of detection be equal to  $\beta_{\min}$  and it follows that

$$c^{(2)}\zeta(\Phi_{(2)}) = c^{(1)}\zeta(\Phi_{(1)}). \quad (\text{I.7})$$

If  $\zeta(\Phi_{(1)}) \neq \zeta(\Phi_{(2)})$ , then  $c^{(1)} \neq c^{(2)}$  and the two sequences of tests would have different hypotheses pairs for each  $n$ ; i.e., the sequence of hypothesis given by  $\theta_{1,n}^{(1)} = c^{(1)}/\sqrt{n}$  would be different from the sequence given by  $\theta_{1,n}^{(2)} = c^{(2)}/\sqrt{n}$ .

In order to compare the two candidate sequence of systems under the same sequence of hypotheses pairs, pick a subsequence of hypotheses  $H_{1,n}^{(1)} := c^{(1)}/\sqrt{d^{(1)}n}$  and a subsequence  $\Phi_{(1*)} := \{\phi_{(1)}^{d^{(1)}n}\}_{n=1}^{\infty}$  for some integer  $d^{(1)}$ ; i.e.,  $\Phi_{(1*)}$  is a sequence of detection systems in which, for each  $n$ , each detection system has  $K = d^{(1)}n$  sensors. Similarly, pick a subsequence of the hypotheses  $H_{1,n}^{(2)}$  and a subsequence  $\Phi_{(2*)}$  for the second system in which, for each  $n$ , each detection system has  $d^{(2)}n$  sensors for some  $d^{(2)}$  integer.

Note that both subsequences will still have their probability of detection converge to  $\beta_{\min}$  since a subsequence of a convergent sequence has the same limit. In other words, the subsequences of probabilities of detection for  $\Phi_{(1)}$  and  $\Phi_{(2)}$  will also converge to (8.8) with their respective efficacies and values for  $c$ .

If the values of  $d^{(1)}$  and  $d^{(2)}$  that define the subsequences are such that

$$\frac{c^{(1)}}{\sqrt{d^{(1)} \cdot n}} = \frac{c^{(2)}}{\sqrt{d^{(2)} \cdot n}} \quad (\text{I.8})$$

then  $H_{1,n}^{(1)} = H_{1,n}^{(2)}$ ; i.e., the comparison will be under the same subsequence of hypothesis as desired.

Thus, if values for  $c^{(1)}$ ,  $c^{(2)}$ ,  $d^{(1)}$ , and  $d^{(2)}$  that satisfy both (I.7) and (I.8) are chosen, then the comparison will be done at the same subsequence of hypothesis and both subsequences of detectors will have the same asymptotic probabilities of detection and false alarm.

The requirement that the comparison be done at the same subsequence of hypothesis and that both subsequences have the same asymptotic detection performance causes the number of sensors used in each subsequence of systems being compared to be different. More precisely, when comparing  $\Phi_{(1)}$  and  $\Phi_{(2)}$ , instead of considering that the number of sensors  $K^{(1)}(n)$  used in each system of  $\Phi_{(1)}$  is equal to the number of sensors  $K^{(2)}(n)$  used in each system of  $\Phi_{(2)}$  for each  $n$ , one considers that  $K^{(1)}(n) = d^{(1)} \cdot n$  and  $K^{(2)}(n) = d^{(2)} \cdot n$ ; and from (I.7) and (I.8), one obtains<sup>1</sup>

$$\frac{K^{(1)}(n)}{K^{(2)}(n)} = \frac{d^{(1)}}{d^{(2)}} = \left( \frac{\zeta(\Phi_{(2)})}{\zeta(\Phi_{(1)})} \right)^2, \quad (\text{I.9})$$

which justifies the definition of the ARE metric as the ratio of the square of the efficacies.

### I.3 Conditions for Quadratic Mean Differentiability and Results for Mixtures of Gaussian and Mixtures of Poisson distributions

This section presents sufficient conditions to ensure that a mixture of distributions is quadratic mean differentiable and further conditions that cause the Fisher information to be proportional to the square of the mean of the mixing random variable. This section also contains propositions that show that these conditions are satisfied by both a mixture of Gaussian and a mixture of Poisson distributions.

Sufficient conditions to ensure that a family of distributions is quadratic mean differentiable at any given  $\theta$  are given in Theorem 12.2.1 of [73]; however, they may be difficult to apply when the distributions in the family are given by a mixture of other distributions. This section presents a proposition that allows one to assert quadratic mean differentiability based only on conditions on the distribution that is being mixed. For such a proposition, Theorem 2.27 of [47] will be used in various instances. To facilitate the understanding, the theorem is stated below.

**Theorem I.1.** [47] *For any  $\varepsilon > 0$  and any space  $\mathcal{X}$ , let  $\mu$  be a measure on  $\mathcal{X}$  with its Borel sigma-algebra. Let  $j(\theta^*, x)$  be a real function that satisfies:  $\forall \theta^* \in [-\varepsilon, \varepsilon], \int_{\mathcal{X}} |j(\theta^*, x)| d\mu(x) < \infty$ . Let  $J(\theta^*) := \int j(\theta^*, x) d\mu(x)$ .*

- *If  $\forall \theta^* \in [-\varepsilon, \varepsilon], \forall x \in \mathcal{X}, |j(\theta^*, x)| \leq c(x)$  for some  $c(x) \in L_1(\mu)$ ; and if  $j(\theta^*, x)$  is a continuous function of  $\theta^*$  for all  $x \in \mathcal{X}$ , then  $J$  is also continuous;*
- *If  $\partial j / \partial \theta^*$  exists and  $\forall \theta^* \in [-\varepsilon, \varepsilon], \forall x \in \mathcal{X}, |\partial j(\theta^*, x) / \partial \theta^*| \leq c(x)$  for some  $c(x) \in L_1(\mu)$ ; then  $J$  is differentiable and  $dJ(\theta^*) / d\theta^* = \int \partial j(\theta^*, x) / \partial \theta^* d\mu(x)$ .*

**Proposition I.2.** *Suppose that  $Z$  is a random variable with realizations in some space  $\mathcal{Z}$  and suppose that  $Z$  depends on a random variable  $X$  with realizations in  $\mathcal{X}$  and on a random variable  $\theta$  with realizations in  $[-\varepsilon, \varepsilon]$  for some*

<sup>1</sup>Integers  $d^{(1)}$  and  $d^{(2)}$  that satisfy (I.9) can be found if the ratio of squared efficacies is rational. If irrational, the ratio provides a lower bound for the number of sensors required in the first sequence of detectors in order to equalize its performance with the second sequence of detectors under a low SNR regime and this lower bound can be as tight as one desires since one can choose a rational number higher than but as close as one desires to the ratio of squared efficacies.



$\varepsilon > 0$ . For any realization  $x$  of  $X$  and any realization  $\theta^*$  of  $\theta$ , let  $P[Z \leq z | \theta = \theta^*, X = x]$  be a c.d.f.; i.e., for any  $x \in \mathcal{X}$  and any  $\theta^* \in [-\varepsilon, \varepsilon]$ ,  $P_{Z|\theta=\theta^*, X=x}$  is a probability measure. Assume that each  $P_{Z|\theta=\theta^*, X=x}$  has a density function  $f_{Z|\theta=\theta^*, X=x}(z)$  with respect to a common measure  $\mu$  and assume that  $P_{Z|\theta=\theta^*}$  has a density function with respect to  $\mu$  given by

$$f_{Z|\theta=\theta^*}(z) = \int f_{Z|\theta=\theta^*, X=x}(z) dP_X(x). \quad (\text{I.10})$$

Let  $g(z, \theta^*, x) := f_{Z|\theta=\theta^*, X=x}(z)$ . If

- For any  $z \in \mathcal{Z}$ , any  $x \in \mathcal{X}$ ,  $g(z, \theta^*, x)$  is a continuous function of  $\theta^*$  on  $[-\varepsilon, \varepsilon]$ ; and

$$\forall \theta^* \in [-\varepsilon, \varepsilon], 0 < g(z, \theta^*, x) < C_1(z) \quad (\text{I.11})$$

for some function  $C_1(z) < \infty$ ;

- For any  $z \in \mathcal{Z}$  and any  $x \in \mathcal{X}$ ,  $\partial g(z, \theta^*, x) / \partial \theta^*$  exists and is a continuous function of  $\theta^*$  for all  $\theta^* \in [-\varepsilon, \varepsilon]$ , and

$$\forall \theta^* \in [-\varepsilon, \varepsilon], \left| \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) \right| < C_2(z) \quad (\text{I.12})$$

for some function  $C_2(z) < \infty$ ;

- There exists a function  $h(z) \in L_1(\mu)$  that satisfies,

$$\forall z \in \mathcal{Z}, \forall \theta^* \in [-\varepsilon, \varepsilon], 0 < \frac{\left( \int \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) dP_X(x) \right)^2}{4 \int g(z, \theta^*, x) dP_X(x)} \leq h(z); \quad (\text{I.13})$$

then the family of densities  $\{f_{Z|\theta=\theta^*}(z) | \theta^* \in (-\varepsilon, \varepsilon)\}$  is quadratic mean differentiable at  $\theta^* = 0$ .

*Proof:* From Theorem 12.2.1 of [73], quadratic mean differentiability at  $\theta^* = 0$  will follow if three conditions are satisfied: (i)  $f_{Z|\theta=\theta^*}^{1/2}(z)$  (or  $f_{Z|\theta=\theta^*}(z)$ ) is an absolutely continuous function of  $\theta^*$  in some neighborhood of 0 for  $\mu$ -almost all  $z$ , (ii)  $\partial f_{Z|\theta=\theta^*}(z) / \partial \theta^*$  exists at  $\theta^* = 0$  for  $\mu$ -almost all  $z$ , and (iii) the Fisher information  $I(\theta^*)$  is finite and continuous in  $\theta^*$  at  $\theta^* = 0$ .

To prove these conditions, it will first be proven that, for any  $z \in \mathcal{Z}$ ,  $f_{Z|\theta=\theta^*}(z)$  is differentiable as a function of  $\theta^*$  and its derivative is a continuous function of  $\theta^* \in [-\varepsilon, \varepsilon]$ .

To prove that  $f_{Z|\theta=\theta^*}(z)$  is differentiable, pick any  $z$  and apply Theorem I.1 with  $j(\theta^*, x) := g(z, \theta^*, x)$ ,  $\mu = P_X$ , and  $J(\theta^*) = f_{Z|\theta=\theta^*}(z)$ , which follows from (I.10). Assumption (I.11) gives that  $\forall \theta^* \in [-\varepsilon, \varepsilon]$ ,  $\int |j(\theta^*, x)| dP_X(x) < \int C_1(z) dP_X(x) = C_1(z) < \infty$  because  $P_X$  is a probability measure. Since it is assumed that  $\partial g(z, \theta^*, x) / \partial \theta^*$  exists and assumption (I.12) gives that its absolute value is bounded by  $C_2(z)$ , let  $c(x) = C_2(z)$  for all  $x$  and  $c(x) \in L_1(P_X)$  because  $P_X$  is a probability measure. Thus, for any  $z \in \mathcal{Z}$ , the conditions for the second conclusion of Theorem I.1 are satisfied, giving that  $f_{Z|\theta=\theta^*}(z)$  as a function of  $\theta^*$  is differentiable and

$$\frac{\partial}{\partial \theta^*} f_{Z|\theta=\theta^*}(z) = \int \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) dP_X(x). \quad (\text{I.14})$$

To prove that  $f_{Z|\theta=\theta^*}(z)$  as a function of  $\theta^*$  has a continuous derivative for all  $\theta^* \in [-\varepsilon, \varepsilon]$  and all  $z \in \mathcal{Z}$ , pick any  $z$  and apply Theorem I.1 with  $j(\theta^*, x) := \partial g(z, \theta^*, x)/\partial \theta^*$  and  $J(\theta^*) = \partial f_{Z|\theta=\theta^*}(z)/\partial \theta^*$ . Note that assumption (I.12) gives that  $\forall \theta^* \in [-\varepsilon, \varepsilon], \forall x \in \mathcal{X}, |j(\theta^*, x)| < c(x) = C_2(z)$  and both  $j(\theta^*, x)$  and  $c(x)$  are in  $L_1(P_X)$  because  $P_X$  is a probability measure. Since it is assumed that  $\partial g(z, \theta^*, x)/\partial \theta^*$  is a continuous function of  $\theta^*$  for any  $x \in \mathcal{X}$  and any  $z \in \mathcal{Z}$ , the first conclusion of Theorem I.1 indicates that  $f_{Z|\theta=\theta^*}(z)$  has continuous derivative with respect to  $\theta^*$ .

These results are now used to prove the condition that  $f_{Z|\theta=\theta^*}(z)$  is absolutely continuous in a neighborhood of 0 (that is a subset of  $[-\varepsilon, \varepsilon]$ ) for any  $z \in \mathcal{Z}$ , which is equivalent to the condition that,  $\forall z \in \mathcal{Z}$ , there exists a function  $a_z(t)$  that satisfies

$$\forall \theta^* \in [-\varepsilon, \varepsilon], f_{Z|\theta=\theta^*}(z) - f_{Z|\theta=-\varepsilon}(z) = \int_{-\varepsilon}^{\theta^*} a_z(t) dt \quad (\text{I.15})$$

and  $\int_{-\varepsilon}^{+\varepsilon} |a_z(t)| dt < \infty$  [47, p. 106]. To show that such a function  $a_z(t)$  exists, use the already proven result that  $f_{Z|\theta=\theta^*}(z)$  is differentiable and has continuous derivative to conclude that  $\partial f_{Z|\theta=\theta^*}(z)/\partial \theta^*$  is Riemann integrable and use the second Fundamental Theorem of Calculus [101], which gives that (I.15) is satisfied when  $a_z(t)$  equals  $\partial f_{Z|\theta=\theta^*}(z)/\partial \theta^*$  applied at  $\theta^* = t$ . Note that  $\int_{-\varepsilon}^{+\varepsilon} |a_z(t)| dt < \infty$  because  $\partial f_{Z|\theta=\theta^*}(z)/\partial \theta^*$  was shown to be continuous on a compact interval, which means that  $|\partial f_{Z|\theta=\theta^*}(z)/\partial \theta^*|$  is bounded and therefore its integral over  $[-\varepsilon, \varepsilon]$  with respect to Lebesgue measure is finite. Since such a function  $a_z(t)$  exists,  $f_{Z|\theta=\theta^*}(z)$  is absolutely continuous on  $[-\varepsilon, \varepsilon]$  for any  $z \in \mathcal{Z}$ .

The second condition to apply Theorem 12.2.1 of [73] also follows from the already proven result that, for any  $z \in \mathcal{Z}$ ,  $f_{Z|\theta=\theta^*}(z)$  has a continuous derivative for all  $\theta^* \in [-\varepsilon, \varepsilon]$  because this means that  $\partial f_{Z|\theta=\theta^*}(z)/\partial \theta^*$  exists at  $\theta^* = 0$  for all  $z \in \mathcal{Z}$ .

To prove the third condition to apply Theorem 12.2.1 of [73]; i.e., to prove that the Fisher information  $I(\theta^*)$  is finite and continuous in  $\theta^*$  at  $\theta^* = 0$ , observe that  $I(\theta^*)$  is given by [73]:

$$I(\theta^*) = 4 \int \eta^2(z, \theta^*) d\mu(z), \quad (\text{I.16})$$

where

$$\eta^2(z, \theta^*) := \frac{\left( \frac{\partial}{\partial \theta^*} f_{Z|\theta=\theta^*}(z) \right)^2}{4 \cdot f_{Z|\theta=\theta^*}(z)}, \quad (\text{I.17})$$

which is well defined since (I.11) implies  $f_{Z|\theta=\theta^*}(z) > 0$ . The proof follows from the use of Theorem I.1 with  $j(\theta^*, z) := \eta^2(z, \theta^*)$ ,  $\mathcal{X} = \mathcal{Z}$ ,  $x = z$ ,  $\mu$  being the measure that dominates the conditional distributions of  $Z$ , and  $I(\theta^*) = 4J(\theta^*)$ . Note that the condition  $\int_{\mathcal{Z}} |j(\theta^*, z)| d\mu(z) < \infty$  is satisfied because (I.14) gives that (I.17) equals the term being bounded in (I.13), which is assumed to be bounded by a function  $h(z) \in L_1(\mu)$ . For the same reason, the condition for the first result of Theorem I.1 is also satisfied. Observe now that the  $j(\theta^*, z)$  is a continuous function of  $\theta^*$  for any  $z$  because  $\eta^2(z, \theta^*)$  is a continuous function of  $\theta^*$  for any  $z \in \mathcal{Z}$  since, as proven above, for any  $z \in \mathcal{Z}$ ,  $f_{Z|\theta=\theta^*}(z)$  is differentiable and therefore continuous at all  $\theta^* \in [-\varepsilon, \varepsilon]$ ; and  $f_{Z|\theta=\theta^*}(z)$  has a continuous derivative for

all  $\theta^* \in [-\varepsilon, \varepsilon]$ . Thus, the first result of Theorem I.1 gives that  $I(\theta^*)$  is continuous for all  $\theta^* \in [-\varepsilon, \varepsilon]$ , which implies that  $I(0) < \infty$ . ■

**Proposition I.3.** *Under the same definitions and conditions of Proposition I.2, if  $g(z, \theta^*, x)$  satisfy the following additional conditions:*

- The function  $g(z, \theta^*, x)$  does not depend on  $x$  when  $\theta^* = 0$ ; i.e.,  $g(z, 0, x) = g(z, 0, 0)$ , and
- There exists a function  $h^*(z)$  such that

$$\forall z \in \mathcal{Z}, \forall x \in \mathcal{X}, \left. \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) \right|_{\theta^*=0} = x \cdot h^*(z), \quad (\text{I.18})$$

then, for a constant  $C < \infty$  that does not depend on the distribution of  $X$ , the Fisher Information at  $\theta^* = 0$  satisfies

$$I(0) = C \cdot E[X]^2. \quad (\text{I.19})$$

*Proof:* The conditions imply that  $\eta^2(z, 0)$  can be written as

$$\eta^2(z, 0) = \frac{\left( \left. \frac{\partial}{\partial \theta^*} [\int g(z, \theta^*, x) dP_X(x)] \right|_{\theta^*=0} \right)^2}{4 \cdot \int g(z, 0, x) dP_X(x)} = \frac{\left( \int \left. \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) \right|_{\theta^*=0} dP_X(x) \right)^2}{4 \cdot \int g(z, 0, x) dP_X(x)} \quad (\text{I.20})$$

$$= \frac{(\int x \cdot h^*(z) dP_X(x))^2}{4 \cdot \int g(z, 0, x) dP_X(x)} = \frac{(h^*(z))^2}{4 \cdot g(z, 0, 0)} \cdot E[X]^2, \quad (\text{I.21})$$

where the first equality of (I.20) follows from the definition (I.17) of  $\eta^2$  and (I.10), and the second equality of (I.20) follows from (I.14).

Thus, the Fisher Information at  $\theta = 0$  can be written as

$$I(0) = \left( \int_0^\infty \frac{(h^*(z))^2}{g(z, 0, 0)} d\mu(z) \right) \cdot E[X]^2, \quad (\text{I.22})$$

and the conclusion follows. ■

The conditions of Propositions I.2 and I.3 are met for two cases of interest, as shown by the following Propositions.

**Proposition I.4.** *Consider the assumptions of Propositions 6.2 or 6.3 and consider the additional Assumption VII of Chapter 5. Consider a sequence  $\{\theta_{1,n}\}_{n=1}^\infty$  that converges to 0. If  $A_i = \theta_{1,n} \cdot \xi(\|L_i - L_e\|)$  and  $W_i$  is Gaussian distributed with mean 0 and variance  $\sigma^2$  for all  $i$ , then the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  with respect to the Lebesgue measure belong to a family of densities that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ ; and the Fisher information satisfies*

$$I_{\text{Gaussian}}(0) = \frac{E[\xi(\|L_i - l_e^*\|)]^2}{\sigma^2} \quad (\text{I.23})$$

for any  $l_e^* \in S_e^*$  as defined in Propositions 6.2 or 6.3.

*Proof:* Observe that Propositions 6.2 or 6.3 indicate that  $\{Z_i\}_{i=1}^K$  are conditionally i.i.d.; thus, it is enough to show that the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  for any chosen  $i$  belong to a family of densities that is quadratic mean differentiable at  $\theta = 0$ . Furthermore, the conditions of Propositions 6.2 or 6.3 imply that the distribution of  $A_i$  is invariant to changes in  $L_e \in S_e^*$  for any chosen  $i$ . More precisely, the conditions of Propositions 6.2 or 6.3 imply the conditions of Lemma 6.1, which in turn implies the conditions of Lemma C.3, which gives that the distribution of  $A_i$  is invariant to changes in  $L_e \in S_e^*$ . Therefore, it is possible to pick any  $l_e^* \in S_e^*$  and let  $P[L_e = l_e^*] = 1$ .

To prove that the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  belong to a family of densities that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ , Proposition I.2 is used to build a family that satisfies this requirement. In order to use this proposition, let the random variables  $Z$  and  $\theta$  of Proposition I.2 refer to  $Z_i$  and  $\theta$  of the measurement model. Let  $\mathcal{X} = \mathbb{R}$ . Let  $X$  refer to  $X_i := \xi(\|L_i - l_e^*\|)$  with  $\mathcal{X} = [0, \xi(0)]$ . Because of the assumptions on  $A_i$  and  $W_i$ ,  $Z_i$  conditioned on  $\{\theta = \theta^*, X_i = x\}$  is distributed according to a Gaussian distribution with variance  $\sigma^2$  and with expected value  $\theta^*x$ ; thus, letting  $\mu$  be Lebesgue measure, the density function  $f_{Z_i|\theta=\theta^*, X_i=x}(z)$  with respect to  $\mu$  exists for all  $\theta^*$  and all  $x$ . Lemma E.1 also gives that  $P_{Z_i|\theta=\theta^*}$  also has a density function with respect to  $\mu$  given by (I.10).

Pick  $\varepsilon = 1$  and build the family of densities  $\{f_{Z_i|\theta=\theta^*}(z) | \theta^* \in (-\varepsilon, \varepsilon)\}$  with  $f_{Z_i|\theta=\theta^*}(z)$  given by (I.10). Note that the conditional densities  $f_{Z_i|\theta=0}(z)$  and  $f_{Z_i|\theta=\theta_{1,n}}(z)$  belong to such a family for large enough  $n$ . To show that this family is quadratic mean differentiable at  $\theta^* = 0$ , it is shown next that  $g(z, \theta^*, x) := f_{Z_i|\theta=\theta^*, X=x}(z)$  satisfies the conditions of Proposition I.2.

The first condition of Proposition I.2 is satisfied because  $g(z, \theta^*, x)$  is the Gaussian probability density with expected value  $\theta^* \cdot x$ , which is a positive function, is a continuous function of  $\theta^*$ , and is bounded by  $C_1(z) := 1/(\sqrt{2\pi} \cdot \sigma)$  for all  $z$ , all  $x$ , and all  $\theta^*$ .

The second condition of Proposition I.2 is satisfied because  $\partial g(z, \theta^*, x)/\partial \theta^* = g(z, \theta^*, x) \cdot (x/\sigma^2) \cdot (z - \theta^*x)$  exists for all  $\theta^*$ , is a continuous function of  $\theta^*$  for all  $\theta^*$ , and  $|\partial g(z, \theta^*, x)/\partial \theta^*| \leq C_2(z) := \xi(0) \cdot (|z| + \xi(0))/(\sqrt{2\pi} \cdot \sigma^3)$ , where it is noted that  $\xi(0) < \infty$  because of Assumption VII.

To show that the third condition is satisfied, let

$$g_m(z) := \min_{\{\theta^* \in [-1, 1], x \in [0, \xi(0)]\}} g(z, \theta^*, x), \quad (\text{I.24})$$

$$g_M(z) := \max_{\{\theta^* \in [-1, 1], x \in [0, \xi(0)]\}} g(z, \theta^*, x), \quad (\text{I.25})$$

and write

$$\begin{aligned} & \frac{\left( \int \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) dP_X(x) \right)^2}{4 \int g(z, \theta^*, x) dP_X(x)} \leq \frac{\left( \int \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) dP_X(x) \right)^2}{4g_m(z)} \\ &= \frac{\left( \int g(z, \theta^*, x) \cdot (x/\sigma^2) \cdot (z - \theta^*x) dP_X(x) \right)^2}{4g_m(z)} \\ &\leq \frac{\left( \int g(z, \theta^*, x) \cdot (x/\sigma^2) \cdot (|z| + |\theta^*x|) dP_X(x) \right)^2}{4g_m(z)} \end{aligned}$$

$$\leq \frac{g_M^2(z) \cdot (\xi(0)/\sigma^2)^2 \cdot (|z| + \xi(0))^2}{4g_m(z)} = \frac{g_M^2(z)}{g_m(z)} (|z| + \xi(0))^2 \cdot C =: h(z), \quad (\text{I.26})$$

where  $C$  is a positive finite value that does not depend on  $z$ .

It is now claimed that  $h(z) \in L_1(\mu)$ . To see this, it will be shown that  $h(z) \leq h_1(z)\mathbf{1}(z > \xi(0)) + h_2(z)\mathbf{1}(z < -\xi(0)) + h_3(z)\mathbf{1}(z \in [-\xi(0), \xi(0)])$ , in which each term belongs to  $L_1(\mu)$ .

Consider first that  $z > \xi(0)$ . In this case,  $g_m(z) = g(z, -1, \xi(0))$  and  $g_M(z) = g(z, 1, \xi(0))$ . Note also that  $|z| + \xi(0) \leq 2|z|$  and

$$\begin{aligned} \forall z > \xi(0), h(z) &\leq \frac{g^2(z, 1, \xi(0))}{g(z, -1, \xi(0))} \cdot 4|z|^2 \cdot C \\ &= \exp\{2z\xi(0)/\sigma^2\} g(z, 1, \xi(0)) \cdot 4|z|^2 \cdot C \\ &= g(z, 1, 3\xi(0)) \exp\{4\xi(0)^2/\sigma^2\} \cdot 4|z|^2 \cdot C =: h_1(z), \end{aligned} \quad (\text{I.27})$$

and  $h_1(z) \in L_1(\mu)$  because  $\int |h_1| d\mu$  equals to a constant multiplied by the second absolute moment of a Gaussian random variable, which is finite.

Similarly, for  $z < -\xi(0)$ ,  $g_m(z) = g(z, 1, \xi(0))$ ,  $g_M(z) = g(z, -1, \xi(0))$ ,  $|z| + \xi(0) \leq 2|z|$ , and

$$\begin{aligned} \forall z < -\xi(0), h(z) &\leq \frac{g^2(z, -1, \xi(0))}{g(z, +1, \xi(0))} \cdot 4|z|^2 \cdot C \\ &= \exp\{-2z\xi(0)/\sigma^2\} g(z, -1, \xi(0)) \cdot 4|z|^2 \cdot C \\ &= g(z, -1, 3\xi(0)) \exp\{4\xi(0)^2/\sigma^2\} \cdot 4|z|^2 \cdot C =: h_2(z), \end{aligned} \quad (\text{I.28})$$

and  $h_2(z) \in L_1(\mu)$ .

For  $z \in [-\xi(0), \xi(0)]$ ,  $g_m(z) = g(2\xi(0), 0, 0)$ ,  $g_M(z) = g(0, 0, 0)$ ,  $|z| + \xi(0) \leq 2\xi(0)$ , and

$$\forall z \in [-\xi(0), \xi(0)], h(z) \leq \frac{g^2(0, 0, 0)}{g(2\xi(0), 0, 0)} \cdot 4\xi(0)^2 \cdot C =: h_3(z), \quad (\text{I.29})$$

which is a constant finite value and  $h_3(z)\mathbf{1}\{z \in [-\xi(0), \xi(0)]\} \in L_1(\mu)$ .

Since all the conditions for Proposition I.2 are satisfied, it follows that the family of densities  $\{f_{Z_i|\theta=\theta^*}(z)|\theta^* \in (-\varepsilon, \varepsilon)\}$  is quadratic mean differentiable at  $\theta^* = 0$ .

To show that the Fisher information satisfies (I.23), Proposition I.3 is used. The first condition for this proposition is satisfied because  $g(z, \theta^*, x)$  is the Gaussian density function with mean  $\theta^*x$  and, therefore, does not depend on  $x$  when  $\theta^* = 0$ ; i.e.,  $g(z, 0, x) = g(z, 0, 0)$ . The second condition is also satisfied because

$$\begin{aligned} \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) \Big|_{\theta^*=0} &= g(z, \theta^*, x) \cdot \frac{x}{\sigma^2} \cdot (z - \theta^*x) \Big|_{\theta^*=0} \\ &= x \cdot g(z, 0, 0) \cdot z/\sigma^2. \end{aligned}$$

Thus,  $h^*(z) = g(z, 0, 0) \cdot z/\sigma^2$ , and using (I.22), (I.23) follows. ■

**Proposition I.5.** Consider the assumptions of Propositions 6.2 or 6.3 and consider the additional Assumption VII of Chapter 5. Consider a sequence  $\{\theta_{1,n}\}_{n=1}^{\infty}$  that converges to 0. If  $A_i$  conditioned on  $\{L_i = l_i, L_e = l_e\}$  is Poisson distributed with parameter given by  $\theta_{1,n}\xi(\|l_i - l_e\|)$  and  $W_i$  is Poisson distributed with parameter  $\lambda_W > 0$ , then the conditional probability mass functions (p.m.f.s)  $P_{Z_i|\theta=0}(z)$  and  $P_{Z_i|\theta=\theta_{1,n}}(z)$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ ; and the Fisher information satisfies

$$I_{\text{Poisson}}(0) = \frac{E[\xi(\|L_i - l_e^*\|)]^2}{\lambda_W} \quad (\text{I.30})$$

for any  $l_e^* \in S_e^*$  as defined in Propositions 6.2 or 6.3.

*Proof:* As in the proof of Proposition I.4, it is enough to show that the conditional p.m.f.s  $P_{Z_i|\theta=0}(z)$  and  $P_{Z_i|\theta=\theta_{1,n}}(z)$  for any chosen  $i$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  and it is possible to pick any  $l_e^* \in S_e^*$  and let  $P[L_e = l_e^*] = 1$ .

To prove that the conditional p.m.f.s  $P_{Z_i|\theta=0}(z)$  and  $P_{Z_i|\theta=\theta_{1,n}}(z)$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ , Proposition I.2 is used to build a family that satisfies this requirement. In order to use this proposition, let the random variables  $Z$  and  $\theta$  of Proposition I.2 refer to  $Z_i$  and  $\theta$  of the measurement model. Let  $\mathcal{Z} = \mathbb{Z}_+$  and  $\mu$  be counting measure. Let  $X$  refer to  $X_i := \xi(\|L_i - l_e^*\|)$  with  $\mathcal{X} = [0, \xi(0)]$ . Because of the assumptions on  $A_i$  and  $W_i$ ,  $Z_i$  conditioned on  $\{\theta = \theta^*, X_i = x\}$  is distributed according to a Poisson distribution with parameter  $\lambda_W + \theta^*x$ ; thus, the p.m.f.  $P_{Z_i|\theta=\theta^*, X_i=x}(z)$  exists for all  $\theta^*$  and all  $x$  that satisfy  $\theta^*x > -\lambda_W$ . Lemma E.1 also gives that  $P_{Z_i|\theta=\theta^*}$  has p.m.f. given by

$$P_{Z|\theta=\theta^*}(z) = \int P_{Z|\theta=\theta^*, X=x}(z) dP_X(x), \quad (\text{I.31})$$

which is (I.10) for the present setting.

Pick  $\varepsilon = \lambda_W / (2\xi(0))$  and build the family of p.m.f.s  $\{P_{Z_i|\theta=\theta^*}(z) | \theta^* \in (-\varepsilon, \varepsilon)\}$  with  $P_{Z_i|\theta=\theta^*}(z)$  given by (I.31). Note that the conditional p.m.f.s  $P_{Z_i|\theta=0}(z)$  and  $P_{Z_i|\theta=\theta_{1,n}}(z)$  belong to such a family for large enough  $n$ . To show that this family is quadratic mean differentiable at  $\theta^* = 0$ , it is shown next that  $g(z, \theta^*, x) := P_{Z|\theta=\theta^*, X=x}(z)$  satisfies the conditions of Proposition I.2.

The first condition of Proposition I.2 is satisfied because  $g(z, \theta^*, x)$  is the Poisson p.m.f. with parameter  $\lambda_W + \theta^*x$ , which is a positive function for all  $z \in \mathcal{Z}$ , is a continuous function of  $\theta^*$  for all  $\theta^* \in [-\varepsilon, \varepsilon]$  and all  $x \in [0, \xi(0)]$ , and is bounded by  $C_1(z) := 1$  for all  $z \in \mathcal{Z}$ .

The second condition of Proposition I.2 is satisfied because

$$\frac{\partial g(z, \theta^*, x)}{\partial \theta^*} = g(z, \theta^*, x) \cdot x \cdot \left( \frac{z}{\lambda_W + \theta^*x} - 1 \right), \quad (\text{I.32})$$

which exists and is a continuous function of  $\theta^*$  for all  $\theta^* \in [-\varepsilon, \varepsilon]$ . Since  $g(z, \theta^*, x) \leq 1$ ,

$$\left| \frac{\partial g(z, \theta^*, x)}{\partial \theta^*} \right| \leq x \cdot \left( \frac{z}{\lambda_W + \theta^*x} + 1 \right) \quad (\text{I.33})$$

$$\leq \xi(0) \cdot \left( \frac{2z}{\lambda_W} + 1 \right) := C_2(z), \quad (\text{I.34})$$

where  $\lambda_W + \theta^*x \geq \lambda_W - \varepsilon\xi(0) = \lambda_W/2$  was used to reach (I.34). Since  $\xi(0) < \infty$  because of Assumption VII, the second condition of Proposition I.2 is satisfied.

To show that the third condition is satisfied, let

$$g_m(z) = \frac{(\lambda_W/2)^z \cdot \exp\{-3\lambda_W/2\}}{z!}, \quad (\text{I.35})$$

$$g_M(z) = \frac{(3\lambda_W/2)^z \cdot \exp\{-\lambda_W/2\}}{z!}, \quad (\text{I.36})$$

and since  $\theta^* \in [-\varepsilon, \varepsilon]$  with  $\varepsilon = \lambda_W/(2\xi(0))$ , it follows that  $\lambda_W/2 \leq \lambda_W + \theta^*x \leq 3\lambda_W/2$  and it is possible to reach that

$$\forall z \in \mathbf{Z}_+, \forall x \in [0, \xi(0)], \forall \theta^* \in [-\varepsilon, \varepsilon], g_m(z) \leq g(z, \theta^*, x) \leq g_M(z). \quad (\text{I.37})$$

With these bounding functions, write

$$\begin{aligned} & \frac{\left( \int \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) dP_X(x) \right)^2}{4 \int g(z, \theta^*, x) dP_X(x)} \leq \frac{\left( \int \frac{\partial}{\partial \theta^*} g(z, \theta^*, x) dP_X(x) \right)^2}{4g_m(z)} \\ &= \frac{\left( \int g(z, \theta^*, x) \cdot x \cdot (z/(\lambda_W + \theta^*x) - 1) dP_X(x) \right)^2}{4g_m(z)} \\ &\leq \frac{\left( \int g(z, \theta^*, x) \cdot x \cdot (z/(\lambda_W + \theta^*x) + 1) dP_X(x) \right)^2}{4g_m(z)} \\ &\leq \frac{g_M^2(z)}{g_m(z)} \cdot \frac{\xi(0)^2}{4} \cdot \left( \frac{2z}{\lambda_W} + 1 \right)^2 =: h(z). \end{aligned} \quad (\text{I.38})$$

It is now claimed that  $h(z) \in L_1(\mu)$ . To see this, observe that  $g_M(z)/g_m(z) = 3^z \exp\{\lambda_W\}$  and write

$$\begin{aligned} h(z) &= g_M(z) \cdot 3^z \exp\{\lambda_W\} \cdot \left( \frac{2z}{\lambda_W} + 1 \right)^2 \cdot \frac{\xi(0)^2}{4} \\ &= \frac{(9\lambda_W/2)^z \exp\{\lambda_W/2\}}{z!} \cdot \left( \frac{2z}{\lambda_W} + 1 \right)^2 \cdot \frac{\xi(0)^2}{4} \end{aligned} \quad (\text{I.39})$$

$$= g(z, \lambda_W/(2\xi(0)), 7\xi(0)) \cdot \exp\{5\lambda_W\} \cdot \left( \frac{2z}{\lambda_W} + 1 \right)^2 \cdot \frac{\xi(0)^2}{4} \quad (\text{I.40})$$

$$= g(z, \lambda_W/(2\xi(0)), 7\xi(0)) \cdot (C' z^2 + C'' z + C''') \quad (\text{I.41})$$

for some positive constants  $C'$ ,  $C''$ , and  $C'''$ . When computing  $\int |h(z)| d\mu(z)$ , recall that  $g(z, \lambda_W/(2\xi(0)), 7\xi(0))$  is a Poisson p.m.f. with parameter  $7\lambda_W/2$ , which means that  $\int |h(z)| d\mu(z) = C' E[Z^2] + C'' E[Z] + C''' < \infty$  because the first and second moments of a Poisson random variable with parameter  $7\lambda_W/2$  are finite.

Since all the conditions for Proposition I.2 are satisfied, it follows that the family of p.m.f.s  $\{P_{Z_i|\theta=\theta^*}(z) | \theta^* \in (-\varepsilon, \varepsilon)\}$  is quadratic mean differentiable at  $\theta^* = 0$ .

To show that the Fisher information satisfies (I.30), Proposition I.3 is used. The first condition for this proposition is satisfied because  $g(z, \theta^*, x)$  is the Poisson p.m.f. with parameter  $\lambda_W + \theta^*x$  and, therefore, does not depend on  $x$  when

$\theta^* = 0$ ; i.e.,  $g(z, 0, x) = g(z, 0, 0)$ . The second condition is also satisfied because the derivative of  $g(z, \theta^*, x)$ , given by (I.32), applied at  $\theta^* = 0$  equals

$$\left. \frac{\partial}{\partial \theta} g(z, \theta^*, x) \right|_{\theta^*=0} = x \cdot g(z, 0, 0) \cdot \left( \frac{z}{\lambda_W} - 1 \right).$$

■

**Proposition I.6.** *Consider the assumptions of Propositions 6.2 or 6.3 and consider the additional Assumption VII of Chapter 5. Consider a sequence  $\{\theta_{1,n}\}_{n=1}^\infty$  that converges to 0. If  $A_i = \theta_{1,n} \cdot \xi(\|L_i - L_e\|)$ ,  $W_i$  is Gaussian distributed with mean 0 and variance  $\sigma^2$  for all  $i$ , and*

$$U_i = \mathbf{1}\{Z_i > t_s\} \quad (\text{I.42})$$

for some  $t_s$  such that  $P_{U_i|\theta=0}(0) \in (0, 1)$  and  $P_{U_i|\theta=\theta_{1,n}}(0) \in (0, 1)$ ; then the conditional probability mass functions (p.m.f.s)  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ .

*Proof:* As in the proof of Proposition I.4, it is enough to show that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  for any chosen  $i$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  and it is possible to pick any  $l_e^* \in S_e^*$  and let  $P[L_e = l_e^*] = 1$ .

To prove that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ , Proposition I.2 is used to build a family that satisfies this requirement. In order to use this proposition, let the random variables  $Z$  and  $\theta$  of Proposition I.2 refer to  $U_i$  and  $\theta$ . Let  $\mathcal{X} = \{0, 1\}$  and  $\mu$  be counting measure. Let  $X$  refer to  $X_i := \xi(\|L_i - l_e^*\|)$  with  $\mathcal{X} = [0, \xi(0)]$ . Because of the assumptions on  $A_i$  and  $W_i$ , as argued in Proposition I.4, the conditional distribution  $P_{Z_i|\theta=\theta^*, X_i=x}$  exists for all  $\theta^*$  and all  $x$ ; and, therefore, so does the conditional p.m.f.  $P_{U_i|\theta=\theta^*, X_i=x}$  since  $U_i$  is a function of  $Z_i$ . It also follows that  $P_{U_i|\theta=\theta^*}$  satisfies (I.10); i.e.,

$$P_{U_i|\theta=\theta^*}(u) = \int P_{U_i|\theta=\theta^*, X_i=x}(u) dP_{X_i}(x). \quad (\text{I.43})$$

Pick  $\varepsilon = 1$  and build the family of p.m.f.s  $\{P_{U_i|\theta=\theta^*}(u) | \theta^* \in (-\varepsilon, \varepsilon)\}$  with  $P_{U_i|\theta=\theta^*}(u)$  given by (I.43). Note that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  belong to such a family for large enough  $n$ . To show that this family is quadratic mean differentiable at  $\theta^* = 0$ , it is shown next that  $g(u, \theta^*, x) := P_{U_i|\theta=\theta^*, X=x}(u)$  satisfies the conditions of Proposition I.2. For that, observe that  $u \in \{0, 1\}$  and

$$g(u, \theta^*, x) = u \cdot P[Z_i > t_s | \theta = \theta^*, X = x] + (1 - u) \cdot (1 - P[Z_i > t_s | \theta = \theta^*, X = x]) \quad (\text{I.44})$$

$$= P[\theta^* x + W_i > t_s] \cdot (2u - 1) + (1 - u) = Q\left(\frac{t_s - \theta^* x}{\sigma}\right) \cdot (2u - 1) + (1 - u), \quad (\text{I.45})$$

where  $Q(y)$  is the complementary c.d.f. of a Gaussian random variable with zero mean and unit variance.

To see that the first condition of Proposition I.2 is satisfied, observe in (I.45) that  $g(u, \theta^*, x)$  is continuous as a function of  $\theta^*$  because  $Q$  is a continuous function of its argument, which is a continuous function of  $\theta^*$ . Note also



that  $0 < g(u, \theta^*, x) < 1 =: C_1(u)$  for any  $u \in \{0, 1\}$  because, when  $u = 0$ ,  $g(u, \theta^*, x) = 1 - Q((t_s - \theta^*x)/\sigma) > 0$ ; and when  $u = 1$ ,  $g(u, \theta^*, x) = Q((t_s - \theta^*x)/\sigma) > 0$ .

To see that the second condition of Proposition I.2 is satisfied, obtain that

$$\frac{\partial}{\partial \theta^*} g(u, \theta^*, x) = f_{Z_i|\theta=\theta^*, X=x}(t_s) \cdot x \cdot (2u - 1), \quad (\text{I.46})$$

where  $f_{Z_i|\theta=\theta^*, X=x}(z)$  is the p.d.f. of a Gaussian random variable with expected value  $\theta^*x$  and variance  $\sigma^2$ . Observe that, for any  $u \in \{0, 1\}$  and any  $x$ , the derivative (I.46) exists and is continuous in  $\theta^*$  for all  $\theta^*$ . Furthermore,  $|\partial g(u, \theta^*, x)/\partial \theta^*| < f_{Z_i|\theta=0, X=0}(0) \cdot \xi(0) =: C_2(z) < \infty$  because the function  $\xi$  that defines  $X$  is bounded as per Assumption VII.

To see that the third condition of Proposition I.2 is satisfied, observe that

$$\frac{\left( \int \frac{\partial}{\partial \theta^*} g(u, \theta^*, x) dP_X(x) \right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} = \frac{\left( \int f_{Z_i|\theta=\theta^*, X=x}(t_s) \cdot x \cdot (2u - 1) dP_X(x) \right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} \quad (\text{I.47})$$

$$= \frac{\left( \int f_{Z_i|\theta=\theta^*, X=x}(t_s) \cdot x dP_X(x) \right)^2}{4 \int g(u, \theta^*, x) dP_X(x)}, \quad (\text{I.48})$$

where the term  $(2u - 1)$  of the numerator of (I.47) can be taken out of the integral and, when squared, equals 1. Since  $f_{Z_i|\theta=\theta^*, X=x}(t_s) > 0$  for any  $x$ , and  $P_X$  does not put all the mass in  $X = 0$  because  $L_i$  is uniformly distributed in  $S_s$ , (I.48) is greater than 0. To show that there exists a  $h(u) \in L_1(\mu)$  that satisfies the third condition of Proposition I.2, recall that it was already shown that  $\forall u \in \{0, 1\}, \forall \theta^* \in [-1, 1], x \in [0, \xi(0)], g(u, \theta^*, x) > 0$ , which means that the denominator of (I.48) is greater than 0, and write

$$\frac{\left( \int f_{Z_i|\theta=\theta^*, X=x}(t_s) \cdot x dP_X(x) \right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} \leq \frac{\left( f_{Z_i|\theta=0, X=0}(0) \cdot \xi(0) \right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} =: h(u), \quad (\text{I.49})$$

and let  $h(u) = 0$  for any other  $u$ . Since  $h$  is nonzero and finite in only a finite number of values for  $u$  and  $\mu$  is counting measure,  $h \in L_1(\mu)$ .

Since all the conditions for Proposition I.2 are satisfied, it follows that the family of p.m.f.s  $\{P_{U_i|\theta=\theta^*}(u) | \theta^* \in (-\varepsilon, \varepsilon)\}$  is quadratic mean differentiable at  $\theta^* = 0$ . ■

**Proposition I.7.** *Consider the assumptions of Propositions 6.2 or 6.3 and consider the additional Assumption VII of Chapter 5. Consider a sequence  $\{\theta_{1,n}\}_{n=1}^\infty$  that converges to 0. If  $A_i$  conditioned on  $\{L_i = l_i, L_e = l_e\}$  is Poisson distributed with parameter given by  $\theta_{1,n}\xi(\|l_i - l_e\|)$ ,  $W_i$  is Poisson distributed with parameter  $\lambda_W > 0$ , and*

$$U_i = \mathbf{1}\{Z_i > t_s\} \quad (\text{I.50})$$

*for some  $t_s$  such that  $P_{U_i|\theta=0}(0) \in (0, 1)$  and  $P_{U_i|\theta=\theta_{1,n}}(0) \in (0, 1)$ ; then the conditional probability mass functions (p.m.f.s)  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ .*

*Proof:* As in the proof of Proposition I.6, it is enough to show that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  for any chosen  $i$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  and it is possible to pick any  $l_e^* \in S_e^*$  and let  $P[L_e = l_e^*] = 1$ .

To prove that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  belong to a family of p.m.f.s that is quadratic mean differentiable at  $\theta = 0$  for large enough  $n$ , Proposition I.2 is used to build a family that satisfies this requirement. In order to use this proposition, let the random variables  $Z$  and  $\theta$  of Proposition I.2 refer to  $U_i$  and  $\theta$ . Let  $\mathcal{X} = \{0, 1\}$  and  $\mu$  be counting measure. Let  $X$  refer to  $X_i := \xi(\|L_i - l_e^*\|)$  with  $\mathcal{X} = [0, \xi(0)]$ . Because of the assumptions on  $A_i$  and  $W_i$ , as argued in Proposition I.5, the conditional distribution  $P_{Z_i|\theta=\theta^*, X_i=x}$  exists for all  $\theta^*$  and all  $x$  that satisfies  $\theta^*x > -\lambda_W$ ; and, therefore, so does the conditional p.m.f.  $P_{U_i|\theta=\theta^*, X_i=x}$  since  $U_i$  is a function of  $Z_i$ . It also follows that  $P_{U_i|\theta=\theta^*}$  satisfies (I.10); i.e., it can be written as in (I.43).

Pick  $\varepsilon = \lambda_W / (2\xi(0))$  and build the family of p.m.f.s  $\{P_{U_i|\theta=\theta^*}(u) | \theta^* \in (-\varepsilon, \varepsilon)\}$  with  $P_{U_i|\theta=\theta^*}(u)$  given by (I.43). Note that the conditional p.m.f.s  $P_{U_i|\theta=0}(u)$  and  $P_{U_i|\theta=\theta_{1,n}}(u)$  belong to such a family for large enough  $n$ . To show that this family is quadratic mean differentiable at  $\theta^* = 0$ , it is shown next that  $g(u, \theta^*, x) := P_{U_i|\theta=\theta^*, X=x}(u)$  satisfies the conditions of Proposition I.2. For that, observe first that  $u \in \{0, 1\}$  and

$$g(u, \theta^*, x) = u \cdot P[Z_i > t_s | \theta = \theta^*, X = x] + (1 - u) \cdot (1 - P[Z_i > t_s | \theta = \theta^*, X = x]) \quad (\text{I.51})$$

$$\begin{aligned} &= P[Z_i > t_s | \theta = \theta^*, X = x] \cdot (2u - 1) + (1 - u) \\ &= \left( 1 - \sum_{j=0}^{\lfloor t_s \rfloor} \frac{(\lambda_W + \theta^* x)^j \exp\{-(\lambda_W + \theta^* x)\}}{j!} \right) \cdot (2u - 1) + (1 - u). \end{aligned} \quad (\text{I.52})$$

To see that the first condition of Proposition I.2 is satisfied, observe in (I.52) that  $g(u, \theta^*, x)$  is continuous as a function of  $\theta^*$ . Note also that  $0 < g(u, \theta^*, x) < 1 =: C_1(u)$  for any  $u \in \{0, 1\}$ . To see this, recall that  $t_s$  is assumed to be such that  $P_{U_i|\theta=0}(0) \in (0, 1)$  and  $P_{U_i|\theta=\theta_{1,n}}(0) \in (0, 1)$ , which means that  $t_s \geq 0$  and  $g(0, \theta^*, x) = 1 - P[Z_i > t_s | \theta = \theta^*, X = x] \in (0, 1)$  and  $g(1, \theta^*, x) = 1 - g(0, \theta^*, x) > 0$ .

To see that the second condition of Proposition I.2 is satisfied, obtain that

$$\begin{aligned} \frac{\partial}{\partial \theta^*} g(u, \theta^*, x) &= \left( - \sum_{j=0}^{\lfloor t_s \rfloor} \frac{(\lambda_W + \theta^* x)^j \exp\{-(\lambda_W + \theta^* x)\}}{j!} \cdot x \cdot \left( \frac{j}{\lambda_W + \theta^* x} - 1 \right) \right) \cdot (2u - 1), \\ &= \frac{(\lambda_W + \theta^* x)^{\lfloor t_s \rfloor} \exp\{-(\lambda_W + \theta^* x)\}}{(\lfloor t_s \rfloor)!} \cdot x \cdot (2u - 1) \end{aligned} \quad (\text{I.53})$$

and observe that, for any  $u \in \{0, 1\}$  and any  $x \in [0, \xi(0)]$ , the derivative (I.53) exists and is continuous in  $\theta^*$  for all  $\theta^* \in [-\varepsilon, \varepsilon]$  since  $\varepsilon = \lambda_W / (2\xi(0))$ . Furthermore,

$$\left| \frac{\partial}{\partial \theta^*} g(u, \theta^*, x) \right| \leq \frac{(2\lambda_W)^{\lfloor t_s \rfloor}}{(\lfloor t_s \rfloor)!} \cdot \xi(0) =: C_2(u) < \infty, \quad (\text{I.54})$$

where (I.54) follows because  $|2u - 1| = 1$  and  $\lambda_W + \theta^* x \leq 2\lambda_W$ .

To see that the third condition of Proposition I.2 is satisfied, observe that

$$\frac{\left(\int \frac{\partial}{\partial \theta^*} g(u, \theta^*, x) dP_X(x)\right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} = \frac{\left(\int (\lambda_W + \theta^* x)^{\lfloor t_s \rfloor} \exp\{-(\lambda_W + \theta^* x)\} / (\lfloor t_s \rfloor)! \cdot x(2u-1) dP_X(x)\right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} \quad (\text{I.55})$$

$$= \frac{\left(\int (\lambda_W + \theta^* x)^{\lfloor t_s \rfloor} \exp\{-(\lambda_W + \theta^* x)\} / (\lfloor t_s \rfloor)! \cdot x dP_X(x)\right)^2}{4 \int g(u, \theta^*, x) dP_X(x)}, \quad (\text{I.56})$$

where the term  $(2u-1)$  of the numerator of (I.55) can be taken out of the integral and, when squared, equals to 1. Since the Poisson p.m.f. is greater than zero for any nonnegative integer, including  $\lfloor t_s \rfloor$ , and  $P_X$  does not put all the mass at  $X=0$  because  $L_i$  is uniformly distributed in  $S_s$ , (I.56) is greater than 0. To show that there exists an  $h(u) \in L_1(\mu)$  that satisfies the third condition of Proposition I.2, proceed from (I.56) and write

$$\frac{\left(\int (\lambda_W + \theta^* x)^{\lfloor t_s \rfloor} \exp\{-(\lambda_W + \theta^* x)\} / (\lfloor t_s \rfloor)! \cdot x dP_X(x)\right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} \leq \frac{\left((2\lambda_W)^{\lfloor t_s \rfloor} / (\lfloor t_s \rfloor)! \cdot \xi(0)\right)^2}{4 \int g(u, \theta^*, x) dP_X(x)} := h(u).$$

To see that  $h(u) \in L_1(\mu)$ , recall that it was shown that  $g(u, \theta^*, x) > 0$  for any  $u \in \{0, 1\}$ , any  $\theta^* \in [-\varepsilon, \varepsilon]$ , and any  $x \in [0, \xi(0)]$ , thus,  $h(0) < \infty$  and  $h(1) < \infty$ . Since  $h(u) = 0$  for any other  $u$  and  $\mu$  is counting measure,  $h \in L_1(\mu)$ .

Since all the conditions for Proposition I.2 are satisfied, it follows that the family of p.m.f.s  $\{P_{U_i|\theta=\theta^*}(u)|\theta^* \in (-\varepsilon, \varepsilon)\}$  is quadratic mean differentiable at  $\theta^* = 0$ . ■

**Proposition I.8.** *If  $f(t)$  is the p.d.f. of a Gaussian random variable with zero mean and variance  $\sigma^2 > 0$ , and if  $F(t)$  is the corresponding c.d.f., then*

$$\arg \max_t \frac{f(t)^2}{F(t)(1-F(t))} = 0. \quad (\text{I.57})$$

*Proof:* It will first be proven that  $t=0$  is a point of local maximum and then it will be proven that no other point of local maximum exists.

To prove that  $t=0$  is a point of local maximum, define  $g(t) := f(t)^2/(F(t)(1-F(t)))$  and observe that  $g(t)$  is continuous, differentiable, its derivative equals

$$g'(t) = g(t) \left[ \frac{f(t)(2F(t)-1)}{F(t)(1-F(t))} - \frac{2t}{\sigma^2} \right], \quad (\text{I.58})$$

and  $g'(0) = 0$ . Let  $h(t) := f(t)(2F(t)-1)/[F(t)(1-F(t))] - 2t/\sigma^2$ . Since  $h(t)$  is also differentiable, it is possible to reach

$$g''(t) = g'(t)h(t) + g(t) \left[ \frac{g(t)}{f(t)} \left( h(t) + \frac{t}{\sigma^2} \right) (2F(t)-1) + 2 \left( g(t) - \frac{1}{\sigma^2} \right) \right] \quad (\text{I.59})$$

and since  $g''(0) = 2(4-2\pi)/(\sigma^2\pi^2) < 0$ ,  $t=0$  is a point of local maximum.

It is now proven that no other point of local maximum exists through contradiction.

Assume that there exists a point of local maximum for some  $t_1 > 0$  and note that it is enough to consider  $t_1 > 0$  since  $g(t) = g(-t)$ . If  $t_1$  is a point of local maximum, then there exist  $0 < t_2 < t_1 < t_3$  in which  $g(t_2) = g(t_3)$ ,  $g'(t_2) > 0$

and  $g'(t_3) < 0$ , which implies  $h(t_2) > 0$  and  $h(t_3) < 0$  because  $g'(t) = g(t)h(t)$  and  $g(t) > 0$  always. Using the definition of  $h(t)$ , it is possible to reach that  $h(t_2) > 0$  is equivalent to  $g(t_2)\sigma^2 > 2t_2f(t_2)/(2F(t_2) - 1)$  and, similarly,  $h(t_3) < 0$  is equivalent to  $g(t_3)\sigma^2 < 2t_3f(t_3)/(2F(t_3) - 1)$ . Since  $g(t_2) = g(t_3)$ , it must follow that

$$\frac{t_2f(t_2)}{2F(t_2) - 1} < \frac{t_3f(t_3)}{2F(t_3) - 1}. \quad (\text{I.60})$$

It is now shown that (I.60) cannot happen when  $t_2 < t_3$  because the function  $tf(t)/(2F(t) - 1)$  is nonincreasing. To see this, simply compute

$$\frac{\partial}{\partial t} \frac{tf(t)}{2F(t) - 1} = \frac{tf(t)}{2F(t) - 1} \left( \frac{1 - t^2/\sigma^2}{t} - \frac{2f(t)}{2F(t) - 1} \right), \quad (\text{I.61})$$

observe that

$$\frac{\partial}{\partial t} \frac{tf(t)}{2F(t) - 1} \leq 0 \Leftrightarrow m(t) := (1 - t^2/\sigma^2)(2F(t) - 1) - 2tf(t) \leq 0, \quad (\text{I.62})$$

and note that  $m(0) = 0$  and  $m'(t) = -2t(2F(t) - 1)/\sigma^2 < 0$ , which means that  $m(t) \leq 0$  for all  $t \geq 0$ . ■

## Appendix J: Computing the ARE Metric

In this appendix, formulas for  $E[\xi(\|L_i - (0,0)\|)]$  and  $E[\xi(\|L_i - (0, r_M)\|)]$  are provided when either  $S_s = S_e$  or  $S_s = c \cdot S_e$  and  $L_i$  is uniformly distributed in  $S_s$  for some cases of interest. Formulas are provided only for the case  $c \geq 1$ . These formulas are used as the starting point for some of the propositions and to obtain the numerical examples of Sections 8.2 and 9.4.

The superscripts (1) and (2) are used to differentiate between the deployment options 1 and 2 described in Section 9.2. For the developments of Section 8.2,  $E[\xi(\|L_i^{(1)} - (0,0)\|)]$  and  $E[\xi(\|L_i^{(1)} - (0, r_M)\|)]$  are used since the developments of Section 8.2 consider the comparison of systems under the same deployment option 1.

For the computations of  $E[\xi(\|L_i^{(1)} - (0,0)\|)]$ , consider the original  $S_e$  and  $S_s$ , which are centered in the origin; however, in order to facilitate the computation of  $E[\xi(\|L_i^{(j)} - (0, r_M)\|)]$  for either  $j = 1$  or  $j = 2$ , translate  $S_s$  by  $(0, -r_M)$  and rotate it by  $\pi$ . These operations cause the point  $(0, r_M)$  to be translated to the origin and  $S_e \subset \{(x, y) : y \geq 0\}$ . Apply the same operations in the distribution of  $L_i^{(j)}$ . Note that these transformations do not change the result of  $E[\xi(\|L_i^{(j)} - (0, r_M)\|)]$  because  $L_i$  is uniformly distributed and Lebesgue measure is invariant to translations and rotations. Let  $S_s^{(*)}$  refer to the translated and rotated  $S_s$ ; i.e.,  $S_s^* = \{(x, -(y - r_M)) : (x, y) \in S_s\}$ , and let  $S_s^{(*)1}$  and  $S_s^{(*)2}$  denote the deployment regions for options 1 and 2 respectively.

### J.1 When the Region of Interest is a Disk

When  $S_e = B_{r_M}(0)$ ,

$$E[\xi(\|L_i^{(1)} - (0,0)\|)] = \frac{2\pi}{\pi r_M^2} \int_0^{r_M} \xi(r) \cdot r dr, \quad (\text{J.1})$$

$$E[\xi(\|L_i^{(1)} - (0, r_M)\|)] = \frac{1}{\pi r_M^2} \int_0^\pi \int_0^{2r_M \sin \omega} \xi(r) \cdot r dr d\omega, \quad (\text{J.2})$$

$$E[\xi(\|L_i^{(2)} - (0, r_M)\|)] = \frac{1}{\pi(c \cdot r_M)^2} \int_0^{2\pi} \int_0^{r_M(\sin \omega + \sqrt{c^2 - \cos^2 \omega})} \xi(r) \cdot r dr d\omega. \quad (\text{J.3})$$

In order to understand the range of integration for  $r$  in (J.2), note that  $S_s^{(*)1}$  is a ball with radius  $r_M$  and centered at  $(0, r_M)$  in the Cartesian coordinates and the circle that forms the boundary of  $S_s^{(*)1}$  is given in polar coordinates by  $\{(2r_M \sin \omega, \omega) : \omega \in [0, \pi]\}$  [55].<sup>1</sup>

In order to understand (J.3), observe that  $S_s^{(*)2}$  is equal to the set  $\{(x, y) : x^2 + (y - r_M)^2 \leq (cr_M)^2\}$  in Cartesian coordinates, which is equal to the set  $\{(r, \omega) : r^2 - r \cdot (2r_M \sin \omega) + r_M^2 \leq (cr_M)^2\}$  in polar coordinates. Thus, the range of integration for  $r$  is found by solving the equation  $r^2 - r \cdot (2r_M \sin \omega) + r_M^2(1 - c^2) = 0$ . The only positive root of this equation is  $r_M(\sin \omega + \sqrt{c^2 - \cos^2 \omega})$ ; the other root is negative because  $c > 1$  in deployment option 2. Note that the range of integration for  $\omega$  is from 0 to  $2\pi$  because the area of  $S_s^{(*)2} \cap \{(x, y) : y < 0\}$  is positive.

<sup>1</sup>To see this, write any  $(r, \omega) \in \{(2r_M \sin \omega, \omega) : \omega \in [0, \pi]\}$  in Cartesian coordinates and find that its distance from  $(0, r_M)$  equals  $r_M$ .

## J.2 When the Region of Interest is a Convex Regular Polygon

Polar coordinates can also be used to obtain  $E[\xi(\|L_i^{(1)} - (0,0)\|)]$  and  $E[\xi(\|L_i^{(j)} - (0,r_M)\|)]$  when  $S_e$  is a convex regular polygon of  $n$  vertices circumscribed by  $B_{r_M}(0)$  with one of the vertices at  $(0,r_M)$ :

$$E[\xi(\|L_i^{(1)} - (0,0)\|)] = \frac{1}{m(S_e)} \int_0^{2\pi} \int_0^{g_n^{(1)}(\omega)} \xi(r) \cdot r dr d\omega, \quad (\text{J.4})$$

$$E[\xi(\|L_i^{(1)} - (0,r_M)\|)] = \frac{1}{m(S_e)} \int_{\pi/n}^{\pi-\pi/n} \int_0^{h_n^{(1)}(\omega)} \xi(r) \cdot r dr d\omega, \quad (\text{J.5})$$

$$E[\xi(\|L_i^{(2)} - (0,r_M)\|)] = \frac{1}{m(S_s^{(*2)})} \int_{-\pi}^{+\pi} \int_0^{h_n^{(2)}(\omega)} \xi(r) \cdot r dr d\omega, \quad (\text{J.6})$$

where  $m(S_e)$  is the area of  $S_e$ ,  $g_n^{(1)}(\omega)$  determines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega) \in S_e$ ,  $h_n^{(1)}(\omega)$  determines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega) \in S_s^{(*1)}$ , and  $h_n^{(2)}(\omega)$  determines the maximum  $r$  that satisfies  $(r \cos \omega, r \sin \omega) \in S_s^{(*2)}$ . To understand the range of integration over  $\omega$  in (J.5), recall that it is assumed that one of the vertices of the polygon that defines  $S_e$  is located at  $(0,r_M)$ . The translation and rotation operations cause this vertex to be at the origin and the two sides that contain the origin are symmetrical with respect to the line  $x = 0$ . The angle formed by the sides that contain the origin is given by  $\pi(1 - 2/n)$ , which is the internal angle of a regular convex polygon of  $n$  vertices [34], and the range of integration follows. To understand the range of integration over  $\omega$  in (J.6), observe that the area of  $S_s^{(*2)} \cap \{(x,y) : y < 0\}$  is positive.

### J.2.1 Special Case: Triangular Region of Interest

Assume  $S_e$  is a regular triangle circumscribed by  $B_{r_M}(0)$  with a vertex at  $(0,r_M)$ .

It is possible to show that  $m(S_e) = 3r_M^2\sqrt{3}/4$  and

$$E[\xi(\|L_i^{(1)} - (0,0)\|)] = \frac{4}{r_M^2\sqrt{3}} \int_{-5\pi/6}^{-\pi/6} \int_0^{r_M/(-2\sin\omega)} \xi(r) \cdot r dr d\omega \quad (\text{J.7})$$

$$E[\xi(\|L_i^{(1)} - (0,r_M)\|)] = \frac{4}{3r_M^2\sqrt{3}} \int_{\pi/3}^{2\pi/3} \int_0^{3r_M/(2\sin\omega)} \xi(r) \cdot r dr d\omega \quad (\text{J.8})$$

$$E[\xi(\|L_i^{(2)} - (0,r_M)\|)] = \frac{8}{3(cr_M)^2\sqrt{3}} \left[ \int_{-\pi/2}^{\tan^{-1}((2+c)/(\sqrt{3}c))} \int_0^{(c-1)r_M/(\sqrt{3}\cos\omega - \sin\omega)} \xi(r) \cdot r dr d\omega + \int_{\tan^{-1}((2+c)/(\sqrt{3}c))}^{\pi/2} \int_0^{(1+c/2)r_M/\sin\omega} \xi(r) \cdot r dr d\omega \right]. \quad (\text{J.9})$$

To understand (J.7), partition the triangle in three smaller triangles, each one formed by two vertices and the origin. Observe that  $\|L_i^{(1)} - (0,0)\|$  is invariant to rotations by  $\pi/3$ , which means that it is enough to consider the integral of  $\xi(\|L_i^{(1)} - (0,0)\|)$  as  $L_i^{(1)}$  varies in the sector  $(-5\pi/6, -\pi/6)$  of  $S_e$  and multiply the final result by 3. Within this sector, any point in  $S_e$  must satisfy  $r \sin \omega \geq -r_M/2$ , which is equivalent as  $r \leq r_M/(-2\sin(\omega))$ .

To understand (J.8), observe that, in polar coordinates,  $S_s^{(*1)} = \{(r, \omega) : \omega \in (\pi/3, 2\pi/3), r \sin \omega \leq 3r_M/2\}$ .

To understand (J.9), observe that the side of the triangle  $S_s^{(*2)}$  that is contained  $\{(x, y) : x \geq 0\}$  is given by the line segment  $\{(x, \sqrt{3}x - (c-1)r_M) : x \in [0, c\sqrt{3}r_M/2]\}$ , which connects the vertices at  $(0, -(c-1)r_M)$  and  $(cr_M\sqrt{3}/2, r_M(1+c/2))$ . The other side of the triangle that contains the vertex  $(cr_M\sqrt{3}/2, r_M(1+c/2))$  is given by a line segment at  $y = r_M(1+c/2)$ . This means that any point in  $S_s^{(*2)} \cap \{(x, y) : x \geq 0\}$  must satisfy, in polar coordinates,  $\sqrt{3}(r \cos \omega) - (c-1)r_M \leq r \sin \omega \leq r_M(1+c/2)$  for any  $\omega \in [-\pi/2, \pi/2]$ , which is equivalent as having  $r \leq (c-1)r_M/(\sqrt{3} \cos \omega - \sin \omega)$  when  $\omega \in [-\pi/2, \tan^{-1}((2+c)/(\sqrt{3}c))]$  and  $r \leq (1+c/2)r_M/\sin \omega$  when  $\omega \in [\tan^{-1}((2+c)/(\sqrt{3}c)), \pi/2]$ .

## J.2.2 Special Case: Square Region of Interest

Assume  $S_e$  is a square circumscribed by  $B_{r_M}(0)$  with a vertex at  $(0, r_M)$ .

It is possible to show that  $m(S_e) = 2r_M^2$  and

$$E[\xi(\|L_i^{(1)} - (0, 0)\|)] = \frac{2}{r_M^2} \int_0^{\pi/2} \int_0^{r_M/(\cos \omega + \sin \omega)} \xi(r) \cdot r dr d\omega \quad (\text{J.10})$$

$$E[\xi(\|L_i^{(1)} - (0, r_M)\|)] = \frac{1}{r_M^2} \int_{\pi/4}^{\pi/2} \int_0^{2r_M/(\cos \omega + \sin \omega)} \xi(r) \cdot r dr d\omega. \quad (\text{J.11})$$

$$E[\xi(\|L_i^{(2)} - (0, r_M)\|)] = \frac{1}{r_M^2} \left[ \int_{-\pi/2}^{\tan^{-1}(1/c)} \int_0^{r_M(c-1)/(\cos \omega - \sin \omega)} \xi(r) \cdot r dr d\omega + \int_{\tan^{-1}(1/c)}^{\pi/2} \int_0^{r_M(c+1)/(\cos \omega + \sin \omega)} \xi(r) \cdot r dr d\omega \right]. \quad (\text{J.12})$$

To understand (J.10), observe that  $E[\xi(\|L_i^{(1)} - (0, 0)\|)]$  is four times the integral of  $\xi(\|L_i^{(1)} - (0, 0)\|)$  over the quadrant  $\{(x, y) : x \geq 0, y \geq 0\}$  and, because the square has a vertex at  $(0, r_M)$  and another vertex at  $(r_M, 0)$ , for any  $(x, y) \in S_e$ ,  $y$  must satisfy  $y \leq r_M - x$ ; which is equivalent to  $r \sin \omega \leq r_M - r \cos \omega$  in polar coordinates.

To understand (J.11), observe that  $E[\xi(\|L_i^{(1)} - (0, r_M)\|)]$  is two times the integral of  $\xi(\|L_i^{(1)} - (0, 0)\|)$  over the quadrant  $\{(x, y) : x \geq 0, y \geq 0\}$ . After the translation and rotation operations,  $S_s^{(*1)}$  has a vertex at  $(0, 2r_M)$  and a vertex at  $(r_M, r_M)$  and the line segment that has these 2 vertices as end points define the maximum range of  $r$  when  $\omega$  varies from  $\pi/4$  to  $\pi/2$ ; thus, any  $(x, y) \in S_s^{(*1)}$  must satisfy  $r \sin \omega \leq 2r_M - r \cos \omega$  in polar coordinates.

To understand (J.12), observe that the side of the square  $S_s^{(*2)}$  that connects the vertices at  $(0, -(c-1)r_M)$  and  $(cr_M, r_M)$  is given by the line segment  $\{(x, x - (c-1)r_M) : x \in [0, cr_M]\}$ . The other side of the square that contains the vertex  $(cr_M, r_M)$  is given by the line segment  $\{(x, -x + (c+1)r_M) : x \in [0, cr_M]\}$ . This means that any point in  $S_s^{(*2)} \cap \{(x, y) : x \geq 0\}$  must satisfy, in polar coordinates,  $r \cos \omega - (c-1)r_M \leq r \sin \omega \leq -r \cos \omega + r_M(c+1)$  for any  $\omega \in [-\pi/2, \pi/2]$ , which is equivalent as having  $r \leq (c-1)r_M/(\cos \omega - \sin \omega)$  when  $\omega \in [-\pi/2, \tan^{-1}(1/c)]$  and  $r \leq (c+1)r_M/(\cos \omega + \sin \omega)$  when  $\omega \in [\tan^{-1}(1/c), \pi/2]$ .

## Appendix K: Supporting Propositions for Chapter 9

**Proposition K.1.** Let  $L_i^{(1)}$  be uniformly distributed in  $S_s = B_{r_M}(0)$  and let  $L_i^{(2)}$  be uniformly distributed in  $S_s = cB_{r_M}(0)$  for  $c > 1$ . Assume  $0 < \xi(r) < \infty$  is continuous,  $\xi(r)$  has continuous derivatives of at least the third order,  $\lim_{r \rightarrow \infty} \xi'(r) = 0$ ,  $\xi''(r) > 0$ , and  $\lim_{r \rightarrow \infty} \xi'''(r) = 0$ . If

$$\lim_{r \rightarrow \infty} r^4 \cdot \xi'''(r) = 0 \quad (\text{K.1})$$

then there exists  $r_M^*$  such that, for all  $r_M > r_M^*$ ,

$$E[\xi(\|L_i^{(2)} - l_e\|)] > E[\xi(\|L_i^{(1)} - l_e\|)], \forall l_e \in \partial B_{r_M}(0), \quad (\text{K.2})$$

for some  $c > 1$ .

*Proof:* Given the circular region and given that  $L_i$  is uniformly distributed in both options,  $E[\xi(\|L_i - l_e\|)]$  is invariant to changes in  $l_e \in \partial B_{r_M}(0)$  and pick  $l_e = (0, r_M)$ .

As in Appendix J, in order to facilitate the computation of  $\|L_i - l_e\|$ , translate the center of  $S_e = B_{r_M}(0)$  and  $S_s$  from the origin to  $(0, -r_M)$ , which causes  $l_e$  to be translated to the origin, and rotate by  $\pi$ .

Using (J.2) and (J.3) from Appendix J, define

$$G(\omega, c) := \int_0^{r_M(\sin \omega + \sqrt{c^2 - \cos^2 \omega})} \xi(r) \cdot r dr. \quad (\text{K.3})$$

and it follows that

$$E[\xi(\|L_i^{(1)} - (0, r_M)\|)] = \int_0^\pi \frac{G(\omega, 1)}{\pi r_M^2} d\omega, \quad (\text{K.4})$$

$$E[\xi(\|L_i^{(2)} - (0, r_M)\|)] = \int_0^{2\pi} \frac{G(\omega, c)}{\pi(c \cdot r_M)^2} d\omega. \quad (\text{K.5})$$

Observe that

$$F(c) := \int_0^\pi \frac{G(\omega, c)}{\pi(c \cdot r_M)^2} d\omega \quad (\text{K.6})$$

is a differentiable function of  $c$ , and

$$F'(c) := \frac{d}{dc} F(c) = \int_0^\pi \frac{\partial}{\partial c} \frac{G(\omega, c)}{\pi(c \cdot r_M)^2} d\omega \quad (\text{K.7})$$

for any  $c \in [1, C]$  with  $1 < C < \infty$ . These claims follow from Theorem I.1 when one considers  $j : (0, \pi) \times [1, 2] \rightarrow \mathbb{R}$  given by  $j(\omega, c) := G(\omega, c)/(\pi c^2 r_M^2)$ . The assumptions that  $\xi(r)$  is a continuous function and  $\xi(0) < \infty$  are sufficient to satisfy the conditions of this theorem.



The strategy of this proof is to show that there exists  $r_M^*$  such that, if  $r_M > r_M^*$ , then  $F'(1) > 0$ , which implies the existence of small enough values of  $c > 1$  for which

$$\begin{aligned} E[\xi(\|L_i^{(1)} - (0, r_M)\|)] &= F(1) < \int_0^\pi \frac{G(\omega, c)}{\pi(c \cdot r_M)^2} d\omega \\ &< \int_0^{2\pi} \frac{G(\omega, c)}{\pi(c \cdot r_M)^2} d\omega = E[\xi(\|L_i^{(2)} - (0, r_M)\|)]. \end{aligned} \quad (\text{K.8})$$

Compute the derivative in (K.7), and find that  $F'(1) > 0$  is equivalent to

$$\int_0^\pi G(\omega, 1) d\omega < \int_0^\pi \xi(2r_M \sin \omega) \cdot r_M^2 d\omega \quad (\text{K.9})$$

Use integration by parts in the left-hand side, observe that  $G(\pi, 1) = 0$ , obtain that

$$\begin{aligned} \int_0^\pi 1 \cdot G(\omega, 1) d\omega &= \pi \cdot G(\pi, 1) - 0 \cdot G(0, 1) - \int_0^\pi \omega \frac{\partial}{\partial \omega} G(\omega, 1) d\omega \\ &= - \int_0^\pi \omega \cdot 2r_M^2 \sin(2\omega) \cdot \xi(2r_M \sin \omega) d\omega, \end{aligned} \quad (\text{K.10})$$

and note that (K.9) is equivalent to

$$\int_0^\pi \xi(2r_M \sin \omega) (2\omega \sin(2\omega) + 1) d\omega > 0. \quad (\text{K.11})$$

Use integration by parts and reach that (K.11) is equivalent to

$$\int_0^\pi \xi'(2r_M \sin \omega) \cdot \cos \omega \cdot \left( 2\omega \sin^2 \omega + \frac{1}{2} \sin 2\omega \right) d\omega < 0, \quad (\text{K.12})$$

where  $\xi'(2r_M \sin \omega)$  indicates the derivative of  $\xi(r)$  with respect to  $r$  applied at  $r = 2r_M \sin \omega$ .

Observe that  $\xi'(2r_M \sin \omega) = \xi'(2r_M \sin(\pi - \omega))$  and decompose (K.12) as

$$\begin{aligned} &\int_0^{\pi/2} \xi'(2r_M \sin \omega) \cdot \cos \omega \cdot \left( 2\omega \sin^2 \omega + \frac{1}{2} \sin 2\omega \right) d\omega \\ &+ \int_{\pi/2}^\pi \xi'(2r_M \sin(\pi - \omega)) \cdot \cos \omega \cdot \left( 2\omega \sin^2 \omega + \frac{1}{2} \sin 2\omega \right) d\omega < 0. \end{aligned} \quad (\text{K.13})$$

Change the integration variable to  $v = \pi - \omega$  in the second term of the left-hand side of (K.13), and reach

$$\begin{aligned} &\int_{\pi/2}^\pi \xi'(2r_M \sin(\pi - \omega)) \cdot \cos \omega \cdot \left( 2\omega \sin^2 \omega + \frac{1}{2} \sin 2\omega \right) d\omega \\ &= \int_0^{\pi/2} \xi'(2r_M \sin v) \cdot (-\cos v) \cdot \left[ 2(\pi - v) \sin^2 v - \frac{1}{2} \sin 2v \right] dv. \end{aligned} \quad (\text{K.14})$$

Use (K.14) in (K.13), observe that the two terms of the left-hand side become under the same integration interval, and reach that (K.12) is equivalent to

$$\int_0^{\pi/2} \xi'(2r_M \sin \omega) \cdot h(\omega) d\omega < 0, \quad (\text{K.15})$$

where  $h(\omega) := \sin(2\omega) \cdot ((2\omega - \pi) \sin \omega + \cos \omega)$ .

It is possible to show that  $h(\omega) > 0$  for  $\omega \in [0, \omega_1]$  and  $h(\omega) \leq 0$  for  $\omega \in (\omega_1, \pi/2]$ , where  $\omega_1 < \pi/4$ . To see this, let  $h_p(\omega) := (2\omega - \pi) \sin \omega + \cos \omega$  and observe that  $h'_p(\omega) = 0$  if and only if  $\tan \omega = \pi - 2\omega$ , which can only happen at a single value of  $\omega \in (0, \pi/2)$ . Since  $h_p(0) > 0$ ,  $h_p(\pi/4) < 0$ ,  $h_p(\pi/2) = 0$ , and  $h_p$  is continuous, there is only one point in which  $h'_p(\omega) = 0$ ; thus, there cannot be any point with  $h_p(\omega) > 0$  when  $\omega \geq \pi/4$ .

Since  $\xi''(r) > 0$ ,  $\xi'(2r_M \sin \omega_1) < \xi'(2r_M \sin \omega)$  for  $\forall \omega \in (\omega_1, \pi/2]$ , and reach that

$$\int_0^{\omega_1} \xi'(2r_M \sin \omega) h(\omega) d\omega + \int_{\omega_1}^{\pi/2} \xi'(2r_M \sin \omega_1) h(\omega) d\omega < 0 \quad (\text{K.16})$$

implies (K.15).

The goal now is to show that there exists a large enough  $r_M^*$  such that (K.16) is true for all  $r_M > r_M^*$ . For this, let  $H(\omega) := \cos(\omega)/2 - 5\cos(3\omega)/18 - \omega \sin(3\omega)/3 - \pi \sin(\omega)/2 + \omega \sin(\omega) + \pi \sin(3\omega)/6$ , which is the primitive of  $h(\omega)$ . Since  $\xi''(r) > 0$ , if  $1/r_M < \omega_1$ , it is possible to reach that

$$\begin{aligned} \int_0^{\omega_1} \xi'(2r_M \sin \omega) h(\omega) d\omega &= \int_0^{1/r_M} \xi'(2r_M \sin \omega) h(\omega) d\omega + \int_{1/r_M}^{\omega_1} \xi'(2r_M \sin \omega) h(\omega) d\omega \\ &< \xi'(2r_M \sin(1/r_M)) [H(1/r_M) - H(0)] + \xi'(2r_M \sin \omega_1) [H(\omega_1) - H(1/r_M)]. \end{aligned} \quad (\text{K.17})$$

Use the bound of (K.17) in (K.16), combine terms, observe that  $H(\pi/2) = 0$ , and rearrange terms to obtain that, if

$$\frac{\xi'(2r_M \sin \omega_1)}{H(1/r_M) - H(0)} > \frac{\xi'(2r_M \sin(1/r_M))}{H(1/r_M)}, \quad (\text{K.18})$$

then (K.16) follows.

As  $r_M \rightarrow \infty$ , the right-hand side of (K.18) converges to the finite value  $\xi'(2)/H(0) < 0$ , and the left-hand side of (K.18) converges to 0 because of (K.1). To see this, use the L'Hospital's rule [101, p. 109] and reach that

$$\lim_{r_M \rightarrow \infty} \frac{\xi'(2r_M \sin \omega_1)}{H(1/r_M) - H(0)} = \lim_{r_M \rightarrow \infty} \frac{r_M^4 \xi'''(2r_M \sin \omega_1) 4 \sin^2 \omega_1}{h'(1/r_M) + 2r_M h(1/r_M)} = 0. \quad (\text{K.19})$$

Therefore, there exists an  $r_M^*$  large enough such that (K.18) is true for all  $r_M > r_M^*$  and the conclusion of the theorem follows. ■

## K.1 Special Case: $\xi(r) = A_{\max}/(1+r)$

**Proposition K.2.** Let  $L_i^{(1)}$  be uniformly distributed in  $S_s = B_{r_M}(0)$  and let  $L_i^{(2)}$  be uniformly distributed in  $S_s = cB_{r_M}(0)$  for  $c > 1$ . If  $\xi(r) = A_{\max}/(1+r)$ , for any  $A_{\max} > 0$ , then there exists  $r_M^*$  such that, for all  $r_M > r_M^*$ ,

$$E[\xi(\|L_i^{(2)} - l_e\|)] > E[\xi(\|L_i^{(1)} - l_e\|)], \forall l_e \in \partial B_{r_M}(0), \quad (\text{K.20})$$

for some  $c > 1$ .

*Proof:* Let  $A_{\max} = 1$  without loss of generality. Follow the same steps of the proof of Proposition K.1 to reach (K.11), which is equivalent to

$$\int_0^\pi \frac{h_0(\omega)}{1 + 2r_M \sin \omega} d\omega > 0, \quad (\text{K.21})$$

in which  $h_0(\omega) := 2\omega \sin(2\omega) + 1$ .

The strategy for this proof is to find an analytically treatable lower bound for the left-hand side of (K.21) and show that the lower bound goes to infinity as  $r_M \rightarrow \infty$ , which implies that (K.21) is true for all  $r_M$  large enough.

It is possible to verify that there exists  $\omega_a$  and  $\omega_b$  such that  $h_0(\omega) \geq 0$  for  $0 \leq \omega \leq \omega_a$  and for  $\omega_b \leq \omega \leq \pi$ ; and  $h_0(\omega) < 0$  for  $\omega_a < \omega < \omega_b$ . Further observe that, since  $\sin 2\omega \geq 0$  for  $\omega \leq \pi/2$  and  $h_0(\pi/2) = 1$ ,  $\pi/2 < \omega_a$ .

For  $\omega_a < \omega < \omega_b$ , use the bound  $h_0(\omega)/(1 + 2r_M \sin \omega) > h_0(\omega)/(2r_M \sin \omega)$ .

For  $0 \leq \omega \leq \omega_a$ , consider  $r_M$  large enough such that  $1/r_M < \pi - \omega_a$ . For  $0 \leq \omega \leq 1/r_M$ ,  $h_0(\omega)/(1 + 2r_M \sin \omega) > 0$ . For  $1/r_M < \omega \leq \omega_a$ , it is claimed that

$$\frac{h_0(\omega)}{1 + 2r_M \sin \omega} > \frac{h_0(\omega)}{2g(r_M)r_M \sin \omega}, \quad (\text{K.22})$$

where  $g(r_M) := 1 + 1/(2r_M \sin(1/r_M))$ . To see this, observe that  $\pi/2 < \omega_a$  implies that  $\pi - \omega_a < \pi/2$  and, for any  $\omega^*, \omega^{**}$  that satisfy  $0 \leq \omega^{**} \leq \pi - \omega_a \leq \omega^* \leq \omega_a$ , it follows that  $\sin \omega^{**} \leq \sin \omega^*$ ; therefore,  $1/r_M < \pi - \omega_a$  implies

$$\begin{aligned} \sin \omega \geq \sin \left( \frac{1}{r_M} \right) &\Leftrightarrow \frac{\sin \omega}{\sin(1/r_M)} + 2r_M \sin \omega \geq 1 + 2r_M \sin \omega \\ \Leftrightarrow \frac{1}{1 + 2r_M \sin \omega} &\geq \frac{1}{\frac{\sin \omega}{\sin(1/r_M)} + 2r_M \sin \omega} = \frac{1}{2r_M \sin \omega \left( 1 + \frac{1}{2r_M \sin(1/r_M)} \right)}. \end{aligned} \quad (\text{K.23})$$

Therefore, if

$$\int_{1/r_M}^{\omega_a} \frac{h_0(\omega)}{2g(r_M)r_M \sin \omega} d\omega + \int_{\omega_a}^{\omega_b} \frac{h_0(\omega)}{2r_M \sin \omega} d\omega > 0 \quad (\text{K.24})$$

then (K.21) follows and the proof is complete.

The primitive of  $h_0(\omega)/\sin \omega$  is given by

$$H_1(\omega) := \log \left( \tan \left( \frac{\omega}{2} \right) \right) + 4 \cos \omega + 4\omega \sin \omega, \quad (\text{K.25})$$

and (K.24) is equivalent to

$$\frac{1}{g(r_M)} [H_1(\omega_a) - H_1(1/r_M)] + H_1(\omega_b) - H_1(\omega_a) > 0. \quad (\text{K.26})$$

To show that the left-hand side of (K.26) grows to infinity, observe that, while both  $H_1(\omega_a)$  and  $H_1(\omega_b)$  have finite values that do not vary as  $r_M \rightarrow \infty$  and  $g(r_M) \rightarrow 1 + 1/2$ ;  $-H_1(1/r_M) \rightarrow \infty$ . ■

## LIST OF REFERENCES

- [1] V. Aalo and R. Viswanathan, "On distributed detection with correlated sensors: two examples," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 25, no. 3, pp. 414–421, May 1989.
- [2] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: a survey," *Computer Networks*, vol. 38, pp. 393–422, 2002.
- [3] G. Anastasi, M. Conti, M. Di Francesco, and A. Passarella, "Energy conservation in wireless sensor networks: a survey," *Ad hoc Networks*, vol. 7, pp. 537–568, 2009.
- [4] S. Appadwedula, V. Veeravalli, and D. Jones, "Robust and locally-optimum decentralized detection with censoring sensors," in *Information Fusion, 2002. Proceedings of the Fifth International Conference on*, vol. 1, 2002, pp. 56–63.
- [5] —, "Energy-efficient detection in sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 23, no. 4, pp. 693–702, April 2005.
- [6] S. Barr, B. Liu, and J. Wang, "Barrier coverage for underwater sensor networks," in *Military Communications Conference, 2008. MILCOM 2008. IEEE*, November 2008, pp. 1–7.
- [7] D. Beale, A. Hume, J. Cage, and P. Williams, "An assessment of multi-static radar remote sensor networks," in *Aerospace Conference, 2004. Proceedings. 2004 IEEE*, vol. 3, March 2004, pp. 2096–2104.
- [8] G. R. Benitz and J. A. Bucklew, "Large deviation rate calculations for nonlinear detectors in gaussian noise," *Information Theory, IEEE Transactions on*, vol. 36, no. 2, pp. 358–371, March 1990.
- [9] R. Blum, "Distributed detection of narrowband signals," *Information Theory, IEEE Transactions on*, vol. 41, no. 2, pp. 519–523, March 1995.
- [10] —, "On the asymptotic relative efficiency of distributed detection schemes," *Information Theory, IEEE Transactions on*, vol. 41, no. 2, pp. 523–527, March 1995.
- [11] —, "Quantization in multisensor random signal detection," *Information Theory, IEEE Transactions on*, vol. 41, no. 1, pp. 204–215, January 1995.
- [12] —, "Locally optimum distributed detection of correlated random signals based on ranks," *Information Theory, IEEE Transactions on*, vol. 42, no. 3, pp. 931–942, May 1996.
- [13] —, "Necessary conditions for optimum distributed sensor detectors under the Neyman–Pearson criterion," *Information Theory, IEEE Transactions on*, vol. 42, no. 3, pp. 990–994, May 1996.
- [14] R. Blum and M. Deans, "Distributed random signal detection with multibit sensor decisions," *Information Theory, IEEE Transactions on*, vol. 44, no. 2, pp. 516–524, March 1998.

- [15] R. Blum and S. Kassam, "Optimal distributed detection of weak signals in dependent sensors," *Information Theory, IEEE Transactions on*, vol. 38, no. 3, pp. 1066–1079, May 1992.
- [16] R. Blum, S. Kassam, and H. Poor, "Distributed detection with multiple sensors: Part II-advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, January 1997.
- [17] R. Blum and B. Sadler, "Energy efficient signal detection in sensor networks using ordered transmissions," *Signal Processing, IEEE Transactions on*, vol. 56, no. 7, pp. 3229–3235, July 2008.
- [18] R. Blum, Z. Xu, and B. Sadler, "Decentralized sensor selection based on sensor observations for energy efficient hypothesis testing," in *Information Sciences and Systems, 2009. CISS 2009. 43rd Annual Conference on*, March 2009, pp. 618–622.
- [19] S. Boyd and L. Vandenbergue, *Convex Optimization*. Cambridge University Press, 2004.
- [20] S. Brennan, A. Mielke, and D. Torney, "Radioactive source detection by sensor networks," *Nuclear Science, IEEE Transactions on*, vol. 52, no. 3, pp. 813–819, June 2005.
- [21] Z. Chair and P. Varshney, "Optimal data fusion in multiple sensor detection systems," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-22, no. 1, pp. 98–101, January 1986.
- [22] J.-F. Chamberland and V. Veeravalli, "Decentralized detection in sensor networks," *Signal Processing, IEEE Transactions on*, vol. 51, no. 2, pp. 407–416, February 2003.
- [23] —, "How dense should a sensor network be for detection with correlated observations?" *Information Theory, IEEE Transactions on*, vol. 52, no. 11, pp. 5099–5106, November 2006.
- [24] —, "Wireless sensors in distributed detection applications," *Signal Processing Magazine, IEEE*, vol. 24, no. 3, pp. 16–25, May 2007.
- [25] —, "Asymptotic results for decentralized detection in power constrained wireless sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1007–1015, August 2004.
- [26] B. Chen, R. Jiang, T. Kasetkasem., and P. Varshney, "Channel aware decision fusion in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 52, no. 12, pp. 3454–3458, December 2004.
- [27] B. Chen and P. Willett, "On the optimality of the likelihood-ratio test for local sensor decision rules in the presence of nonideal channels," *Information Theory, IEEE Transactions on*, vol. 51, no. 2, pp. 693–699, February 2005.
- [28] H. Chen, P. Varshney, and B. Chen, "Conditional dependence in distributed detection: How far can we go?" in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*, June 2009, pp. 664–668.
- [29] —, "A novel framework for distributed detection with dependent observations," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, March 2010, pp. 3526–3529.
- [30] P.-N. Chen and A. Papamarcou, "Likelihood ratio partitions for distributed signal detection in correlated gaussian noise," in *Information Theory, 1995. Proceedings., 1995 IEEE International Symposium on*, September 1995, p. 118.
- [31] Q. Cheng, B. Chen, and P. Varshney, "Detection performance limits for distributed sensor networks in the presence of nonideal channels," *Wireless Communications, IEEE Transactions on*, vol. 5, no. 11, pp. 3034–3038, November 2006.

- [32] S. H. Cho and A. Chandrakasan, "Energy efficient protocols for low duty cycle wireless microsensor networks," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol. 4, 2001, pp. 2041–2044 vol.4.
- [33] C.-Y. Chong and S. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, August 2003.
- [34] H. S. M. Coxeter, *Regular Polytopes. Third Edition.* Dover, 1973.
- [35] A. D'Costa, V. Ramachandran, and A. Sayeed, "Distributed classification of Gaussian space-time sources in wireless sensor networks," *Selected Areas in Communications, IEEE Journal on*, vol. 22, no. 6, pp. 1026–1036, August 2004.
- [36] A. D'Costa and A. Sayeed, "Data versus decision fusion in wireless sensor networks," in *Acoustics, Speech and Signal Processing, 2003. ICASSP 2003. IEEE International Conference on*, vol. IV, 2003, pp. 832–835.
- [37] H. Delic, P. Papantoni-Kazakos, and D. Kazakos, "Fundamental structures and asymptotic performance criteria in decentralized binary hypothesis testing," *Communications, IEEE Transactions on*, vol. 43, no. 1, pp. 32–43, January 1995.
- [38] A. Dogandzic and B. Zhang, "Distributed estimation and detection for sensor networks using hidden markov random field models," *Signal Processing, IEEE Transactions on*, vol. 54, no. 8, pp. 3200–3215, August 2006.
- [39] E. Drakopoulos and C.-C. Lee, "Optimum multisensor fusion of correlated local decisions," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 27, no. 4, pp. 593–606, July 1991.
- [40] P. Dubois, C. Botteron, V. Mitev, C. Menon, P.-A. Farine, P. Dainesie, A. Ionescu, and H. Shea, "Ad hoc wireless sensor networks for exploration of solar-system bodies," *Acta Astronautica*, vol. 64, pp. 626–643, 2009.
- [41] T. Duman and M. Salehi, "Decentralized detection over multiple-access channels," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 34, no. 2, pp. 469–476, April 1998.
- [42] R. Durrett, *Probability: Theory and Examples. Third Edition.* Brooks/Cole, 2005.
- [43] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *Acoustics, Speech and Signal Processing, 2001. ICASSP 2001. IEEE International Conference on*, May 2001, pp. 2033–2036.
- [44] G. Ferrari, M. Martalo, and R. Pagliari, "Decentralized detection in clustered sensor networks," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 47, no. 2, pp. 959–973, April 2011.
- [45] G. Ferrari and R. Pagliari, "Decentralized binary detection with noisy communication links," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 42, no. 4, pp. 1554–1563, October 2006.
- [46] G. Ferrari, R. Pagliari, and M. Martalo, "Decentralized binary detection with non-constant snr profile at the sensors," *Int. J. Sensor Networks*, vol. 4, no. 1/2, pp. 23–36, 2008.
- [47] G. B. Folland, *Real Analysis. Modern Techniques and Their Applications. Second Edition.* John Wiley & Sons, 1999.
- [48] K. Fukuda and T. Uno, "Algorithms for maximizing the volume of intersection of polytopes," in *22nd European Workshop on Computational Geometry*, March 2006, pp. 197–200.

- [49] R. J. Gardner, "The Brunn–Minkowski inequality," *Bulletin of the American Mathematical Society*, vol. 39, no. 3, pp. 355–405, April 2002.
- [50] E. Geraniotis and Y. Chau, "Robust data fusion for multisensor detection systems," *Information Theory, IEEE Transactions on*, vol. 36, no. 6, pp. 1265–1279, November 1990.
- [51] G. H. Golub and C. F. V. Loan, *Matrix Computations. Third Edition*. The John Hopkins University Press, 1996.
- [52] K. Gomadam and S. Jafar, "On the capacity of memoryless relay networks," in *Communications, 2006. ICC '06. IEEE International Conference on*, vol. 4, June 2006, pp. 1580–1585.
- [53] C. H. Gowda and R. Viswanathan, "Robustness of decentralized tests with  $\varepsilon$ -contamination prior," *Information Theory, IEEE Transactions on*, vol. 41, no. 4, pp. 1164–1169, July 1995.
- [54] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. Cambridge Univ. Press, 2006.
- [55] H. L. Guidorizzi, *Um Curso de Calculo - Vol. 2*. LTC - Livros Tecnicos e Cientificos Editora S.A., 1986.
- [56] D. Hall and A. Garga, "Pitfalls in data fusion (and how to avoid them)," in *Information Fusion, 1999. Second International Conference on*, 1999, pp. 429–436.
- [57] M. Hefeeda and M. Bagheri, "Wireless sensor networks for early detection of forest fires," in *Mobile Adhoc and Sensor Systems, 2007. MASS 2007. IEEE International Conference on*, October 2007, pp. 1–6.
- [58] R. Hills, "Sensing for danger," <https://www.llnl.gov/str/JulAug01/Hills.html>, July 2001.
- [59] P. Huber, "A robust version of the probability ratio," *Ann. Math. Statist.*, vol. 36, pp. 1753–1758, 1965.
- [60] P. Huber and V. Strassen, "Minimax tests and the Neyman–Pearson lemma for capacities," *The Annals of Statistics*, vol. 1, no. 2, pp. 251–263, 1973.
- [61] E. Jennings and C. Okino, "On the diameter of sensor networks," in *Aerospace Conference Proceedings, 2002. IEEE*, vol. 3, 2002, pp. 1211–1217.
- [62] R. Jiang and B. Chen, "Fusion of censored decisions in wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 6, pp. 2668–2673, November 2005.
- [63] M. Kam, Q. Zhu, and W. Gray, "Optimal data fusion of correlated local decisions in multiple sensor detection systems," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 28, no. 3, pp. 916–920, July 1992.
- [64] V. Kapnadak, M. Senel, and E. J. Coyle, "Distributed iterative quantization for interference characterization in wireless networks," *Digital Signal Processing*, vol. 22, no. 1, pp. 96–105, January 2012.
- [65] S. M. Kay, *Fundamentals of Statistical Signal Processing : Detection Theory*. Prentice-Hall, 1998.
- [66] W. Knight, R. Pridham, and S. M. Kay, "Digital signal processing for sonar," *Proceedings of the IEEE*, vol. 69, no. 11, pp. 1451–1506, November 1981.
- [67] V. Koch and R. Westphal, "New approach to a multistatic passive radar sensor for air/space defense," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 10, no. 11, pp. 24–32, November 1995.

- [68] M. Kohvakka, M. Kuorilehto, M. Hännikäinen, and T. D. Hämäläinen, "Performance analysis of ieee 802.15.4 and zigbee for large-scale wireless sensor network applications," in *Proceedings of the 3rd ACM international workshop on Performance evaluation of wireless ad hoc, sensor and ubiquitous networks*, ser. PE-WASUN '06, 2006, pp. 48–57.
- [69] M. Kubisch, H. Karl, A. Wolisz, L. Zhong, and J. Rabaey, "Distributed algorithms for transmission power control in wireless sensor networks," in *Wireless Communications and Networking, 2003. WCNC 2003. 2003 IEEE*, vol. 1, March 2003, pp. 558–563.
- [70] S. Kullback, *Information Theory and Statistics*. Dover, 1968.
- [71] J. F. Kurose and K. W. Ross, *Computer Networking: A Top-Down Approach. Forth Edition*. Addison Wesley, 2008.
- [72] G. Lauer and N. Sandell, "Distributed detection with waveform observations: Correlated observation processes," in *American Control Conference, 1982*, June 1982, pp. 812–819.
- [73] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses: Third Edition*. Springer, 2005.
- [74] F. Li and J. Evans, "Design of distributed detection schemes for multiaccess channels," in *Communications Theory Workshop, 2008. AusCTW 2008. Australian*, March 2008, pp. 51–57.
- [75] W. Li and H. Dai, "Distributed detection in wireless sensor networks using a multiple access channel," *Signal Processing, IEEE Transactions on*, vol. 55, no. 3, pp. 822–833, March 2007.
- [76] J. Liang and Q. Liang, "Design and analysis of distributed radar sensor networks," *Parallel and Distributed Systems, IEEE Transactions on*, vol. 22, no. 11, pp. 1926–1933, November 2011.
- [77] X. Lin and R. Blum, "Numerical solutions for optimum distributed detection of known signals in dependent t-distributed noise-the two sensor problem," in *Signals, Systems Computers, 1998. Conference Record of the Thirty-Second Asilomar Conference on*, vol. 1, November 1998, pp. 613–617 vol.1.
- [78] Y. Lin, B. Chen, and L. Tong, "Distributed detection over multiple access channels," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, April 2007, pp. 541–544.
- [79] B. Liu and B. Chen, "Channel-optimized quantizers for decentralized detection in sensor networks," *Information Theory, IEEE Transactions on*, vol. 52, no. 7, pp. 3349–3358, July 2006.
- [80] —, "Decentralized detection in wireless sensor networks with channel fading statistics," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 4, May 2006, pp. IV–IV.
- [81] K. Liu and A. M. Sayeed, "Type-based decentralized detection in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1899–1910, May 2007.
- [82] G. Mergen, V. Naware, and L. Tong, "Asymptotic detection performance of type-based multiple access over multiaccess fading channels," *Signal Processing, IEEE Transactions on*, vol. 55, no. 3, pp. 1081–1092, March 2007.
- [83] R. Min, M. Bhardwaj, S.-H. Cho, N. Ickes, E. Shih, A. Sinha, A. Wang, and A. Chandrakasan, "Energy-centric enabling tecumologies for wireless sensor networks," *Wireless Communications, IEEE*, vol. 9, no. 4, pp. 28–39, August 2002.



- [84] S. Misra, A. Swami, and B. Chen, "Decision fusion in large sensor networks using partially coherent and noncoherent strategies," in *Military Communications Conference, 2007. MILCOM 2007. IEEE*, October 2007, pp. 1–7.
- [85] R. Mudumbai, G. Barriac, and U. Madhow, "On the feasibility of distributed beamforming in wireless networks," *Communications, IEEE Transactions on*, vol. 6, no. 4, pp. 1–10, April 2007.
- [86] R. Mudumbai, B. Wild, U. Madhow, and K. Ramchandran, "Distributed beamforming using 1 bit feedback: from concept to realization," *Allerton Conference on Communication, Control, and Computing*, 2006.
- [87] R. J. Nemzek, J. S. Dreicer, D. C. Torney, and T. T. Warnock, "Distributed detection in a large wireless sensor network," *Nuclear Science, IEEE Transactions on*, vol. 51, no. 4, pp. 1693–1700, August 2004.
- [88] K. Nguyen, T. Alpcan, and T. Basar, "Distributed hypothesis testing with a fusion center: The conditionally dependent case," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*, December 2008, pp. 4164–4169.
- [89] R. Niu, B. Chen, and P. Varshney, "Fusion of decisions transmitted over rayleigh fading channels in wireless sensor networks," *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 1018–1027, March 2006.
- [90] R. Niu, P. K. Varshney, and Q. Cheng, "Distributed detection in a large wireless sensor network," *Information Fusion*, vol. 7, pp. 380–394, 2006.
- [91] N. Pantazis and D. Vergados, "A survey on power control issues in wireless sensor networks," *Communications Surveys Tutorials, IEEE*, vol. 9, no. 4, pp. 86–107, 2007.
- [92] H. V. Poor, *An Introduction to Signal Detection and Estimation. Second Edition*. Springer-Verlag, 1994.
- [93] G. J. Pottie and W. Kaiser, "Wireless integrated network sensors," *Communications of the ACM*, vol. 43, no. 5, pp. 51–58, May 2000.
- [94] Z. Quan, S. Cui, H. V. Poor, and A. H. Sayed, "Collaborative wideband sensing for cognitive radios," *Signal Processing Magazine, IEEE*, vol. 25, no. 6, pp. 60–73, November 2008.
- [95] J. Quiao and R. Blum, "Threshold optimization for distributed order-statistic cfar signal detection," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 32, no. 1, pp. 368–377, January 1996.
- [96] V. Raghunathan, C. Schurgers, S. Park, and M. Srivastava, "Energy-aware wireless microsensor networks," *Signal Processing Magazine, IEEE*, vol. 19, no. 2, pp. 40–50, March 2002.
- [97] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 32, no. 2, pp. 554–568, April 1996.
- [98] N. S. V. Rao, M. Shankar, J.-C. Chin, D. K. Y. Yau, S. Srivathsan, S. S. Iyengar, Y. Yang, and J. C. Hou, "Identification of low-level point radiation sources using a sensor network," in *IPSN '08: Proc. of the 7th Int. Conf. on Inf. Proc. in Sensor Net.*, 2008, pp. 493–504.
- [99] A. Reibman and L. Nolte, "Optimal design and performance of distributed signal detection systems with faults," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 38, no. 10, pp. 1771–1782, October 1990.
- [100] K. Romer and F. Mattern, "The design space of wireless sensor networks," *Wireless Communications, IEEE*, vol. 11, no. 6, pp. 54–61, December 2004.
- [101] W. Rudin, *Principles of Mathematical Analysis*. McGraw-Hill, 1976.

- [102] B. Sadler, "Fundamentals of energy-constrained sensor network systems," *Aerospace and Electronic Systems Magazine, IEEE*, vol. 20, no. 8, pp. 17–35, August 2005.
- [103] H. Soliman, K. Sudan, and A. Mishra, "A smart forest-fire early detection sensory system: Another approach of utilizing wireless sensor and neural networks," in *Sensors, 2010 IEEE*, November 2010, pp. 1900–1904.
- [104] B. Son, Y.-S. Her, and J.-G. Kim, "A design and implementation of forest-fires surveillance system based on wireless sensor networks for South Korea mountains," *International Journal of Computer Science and Network Security*, vol. 6, no. 9B, pp. 124–130, September 2006.
- [105] X. Song, P. Willett, J. Glaz, and S. Zhou, "Active detection with a barrier sensor network using a scan statistic," *Oceanic Engineering, IEEE Journal of*, vol. 37, no. 1, pp. 66–74, January 2012.
- [106] A. Sundaresan, P. Varshney, and N. Rao, "Distributed detection of a nuclear radioactive source using fusion of correlated decisions," in *Information Fusion, 2007 10th Int. Conf. on*, July 2007, pp. 1–7.
- [107] —, "Distributed detection of a nuclear radioactive source based on a hierarchical source model," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, April 2009, pp. 2901–2904.
- [108] Y. Sung, L. Tong, and A. Swami, "Asymptotic locally optimal detector for large-scale sensor networks under the Poisson regime," *Signal Processing, IEEE Transactions on*, vol. 53, no. 6, pp. 2005–2017, June 2005.
- [109] A. Swami, *Wireless sensor networks: signal processing and communications perspectives*. John Wiley and Sons, 2007.
- [110] R. Szewczyk, J. Polastre, A. Mainwaring, and D. Culler, "Lessons from a sensor network expedition," in *Wireless Sensor Networks*, ser. Lecture Notes in Computer Science, H. Karl, A. Wolisz, and A. Willig, Eds. Springer Berlin / Heidelberg, 2004, vol. 2920, pp. 307–322.
- [111] R. Tandra, A. Sahai, and S. Mishra, "What is a spectrum hole and what does it take to recognize one?" *Proceedings of the IEEE*, vol. 97, no. 5, pp. 824–848, May 2009.
- [112] R. Tenney and N. Sandell, "Detection with distributed sensors," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-17, no. 4, pp. 501–510, July 1981.
- [113] S. Thomopoulos and L. Zhang, "Distributed decision fusion in the presence of networking delays and channel errors," *Information Sciences*, vol. 66, pp. 91–118, 1992.
- [114] J. Tsitsiklis, "Extremal properties of likelihood-ratio quantizers," *Communications, IEEE Transactions on*, vol. 41, no. 4, pp. 550–558, April 1993.
- [115] J. Tsitsiklis and M. Athans, "On the complexity of decentralized decision making and detection problems," *Automatic Control, IEEE Transactions on*, vol. 30, no. 5, pp. 440–446, May 1985.
- [116] J. N. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems*, vol. 1, pp. 167–182, 1988.
- [117] J. Tsitsiklis, "Decentralized detection," in *Advances in Statistical Signal Processing, Signal Detection*, H. Poor and J. Thomas, Eds. Greenwich, CT: JAI, 1993, vol. 2.
- [118] Z. Tu and R. Blum, "On the limitations of random sensor placement for distributed signal detection," in *Communications, 2007. ICC '07. IEEE International Conference on*, June 2007, pp. 3195–3200.

- [119] —, “On the limitations of random sensor placement for distributed signal detection,” *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 45, no. 2, pp. 555–563, April 2009.
- [120] J. Unnikrishnan and V. Veeravalli, “Cooperative spectrum sensing and detection for cognitive radio,” in *Global Telecommunications Conf., 2007. GLOBECOM '07. IEEE*, November 2007, pp. 2972–2976.
- [121] A. W. van der Vaart, *Asymptotic Statistics*. Cambridge University Press, 1998.
- [122] P. K. Varshney, *Distributed Detection and Data Fusion*. Springer-Verlag New York, 1996.
- [123] V. Veeravalli, T. Basar, and H. Poor, “Minimax robust decentralized detection,” *Information Theory, IEEE Transactions on*, vol. 40, no. 1, pp. 35–40, January 1994.
- [124] V. Veeravalli and P. K. Varshney, “Distributed inference in wireless sensor networks,” *Philosophical Transactions of the Royal Society-Part A: mathematical, physical, and engineering sciences*, vol. 370, pp. 100–117, January 2012.
- [125] K. E. Viall, “Scatterable sensor networks for network centric warfare applications,” in *Military Communications Conference, 2007. MILCOM 2007. IEEE*, October 2007, pp. 1–7.
- [126] E. Visotsky, S. Kuffner, and R. Peterson, “On collaborative detection of TV transmissions in support of dynamic spectrum sharing,” in *New Frontiers in Dynamic Spectrum Access Networks, 2005. DySPAN 2005. 2005 First IEEE International Symposium on*, November 2005, pp. 338–345.
- [127] T.-Y. Wang and Q. Cheng, “Collaborative event-region and boundary-region detections in wireless sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 6, pp. 2547–2561, June 2008.
- [128] P. Willett and P. Swaszek, “Another paper about the sign-detector,” in *Proceedings of the 27th annual Conference on Information Sciences and Systems*, March 1993, pp. 278–283.
- [129] —, “On the performance degradation from one-bit quantized detection,” *Information Theory, IEEE Transactions on*, vol. 41, no. 6, pp. 1997–2003, November 1995.
- [130] P. Willett, P. Swaszek, and R. Blum, “The good, bad and ugly: distributed detection of a known signal in dependent gaussian noise,” *Signal Processing, IEEE Transactions on*, vol. 48, no. 12, pp. 3266–3279, December 2000.
- [131] P. Willett and D. Warren, “The suboptimality of randomized tests in distributed and quantized detection systems,” *Information Theory, IEEE Transactions on*, vol. 38, no. 2, pp. 355–361, March 1992.
- [132] Q. Yan and R. Blum, “On some unresolved issues in finding optimum distributed detection schemes,” *Signal Processing, IEEE Transactions on*, vol. 48, no. 12, pp. 3280–3288, December 2000.
- [133] —, “Distributed signal detection under the Neyman–Pearson criterion,” *Information Theory, IEEE Transactions on*, vol. 47, no. 4, pp. 1368–1377, May 2001.
- [134] Y. Yang, R. Blum, and B. Sadler, “A distributed and energy-efficient framework for Neyman–Pearson detection of fluctuating signals in large-scale sensor networks,” *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 7, pp. 1149–1158, September 2010.
- [135] S. Yiu and R. Schober, “Non-orthogonal transmission and noncoherent fusion of censored decisions,” in *Communications, 2007. ICC '07. IEEE International Conference on*, June 2007, pp. 980–985.

- [136] M. Younis, M. Youssef, and K. Arisha, "Energy-aware routing in cluster-based sensor networks," in *Modeling, Analysis and Simulation of Computer and Telecommunications Systems, 2002. MASCOTS 2002. Proceedings. 10th IEEE International Symposium on*, 2002, pp. 129–136.
- [137] L. Yu, L. Yuan, G. Qu, and A. Ephremides, "Energy-driven detection scheme with guaranteed accuracy," in *Information Processing in Sensor Networks, 2006. IPSN '06. International Conference on*, April 2006, pp. 284–291.
- [138] S. Zahedi, M. B. Srivastava, and C. Bisdikian, "A framework for quality of information analysis for detection-oriented sensor network deployments," in *IBM Research Report RC24404 (W0711-038)*, November 2007.
- [139] S. Zarrin and T. Lim, "Composite hypothesis testing for cooperative spectrum sensing in cognitive radio," in *Communications, 2009. ICC '09. IEEE International Conference on*, June 2009, pp. 1–5.
- [140] J. Zhang, W. Li, N. Han, and J. Kan, "Forest fire detection system based on a zigbee wireless sensor network," *Frontiers For China*, vol. 3, no. 3, pp. 369–374, 2008.