

Distributed Detection, Localization, and Estimation in Time-Critical Wireless Sensor Networks



UNIVERSITY OF LEEDS

Sami Ahmed Odeh Aldalahmeh

Submitted in accordance with the requirements for the
Degree of Doctor of Philosophy

The University of Leeds
School of Electronic and Electrical Engineering

May 2013

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and other authors to his work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgment.

©2013. The University of Leeds. Sami Ahmed Aldalahmeh.

To my son Ahmad

Acknowledgment

First and foremost, I praise Allah the Almighty the Merciful for his blessings and guidance throughout my journey in the PhD and in my life. No words of gratitude are sufficient to acknowledge his favor on me, no deeds of grace are enough to thank him. So I stand humbled in front of him and ask for his forgiveness, mercy, and aid in the days to come.

I am grateful to Professor Mounir Ghogho who was my academic supervisor during my doctoral studies. His guidance and support were invaluable to my research. Working with Prof. Ghogho was a rich experience in which I learned so much from him, especially how to challenge myself into becoming better researcher academically and personally. I also thank him sincerely for being patient with my shortcomings.

I am thankful to Dr. McLernon for his great help, assistance, and professionalism that I needed in critical times during my research. I also thank Dr. Freear for helping and allowing me to do the engineering I enjoy. I am also grateful to Dr. Marco Cardenas for his help in reviewing my thesis.

I am in an ever unpayable debt to my mother and father for their love, care, and support. I would not become the man I am today without my father who traded his golden days for our sake or my mother who sold her nights for our care. Nothing I say can even begin to describe their favor on me. I hope that I can make up for all the time I was away. I thank my wife for being there for me in the tough times, making them easier, and the good times, making them even better. I also thank my brothers and their cheerful chats and my little sisters for their adorable talks about their studies. I am also in debt to my dear uncle Ghazi for being the man I wanted to be like in every way.

Finally, I extend my dearest thanks to my friends (whom are now all doctors) that I met in my time at the PhD, Amar, FethAllah, Waleed, and Aref for the wonderful time that we had together especially in Willow Close. The marginal talks at dawn with Amar, FethAllah's funny stories, Waleed's words of wisdom, and Aref that waked us up for the dawn prayer and his famous curry dishes. I also thank my dear friend Yossif for his great help when I needed it the most. They were there for me in the time of need and helped me keep going on through the toughest times.

Sami Ahmed Aldalahmeh, Amman.

Abstract

In this thesis the problem of distributed detection, localization, and estimation (DDLE) of a stationary target in a fusion center (FC) based wireless sensor network (WSN) is considered. The communication process is subject to time-critical operation, restricted power and bandwidth (BW) resources operating over a shared communication channel suffering from Rayleigh fading and phase noise. A novel algorithm is proposed to solve the DDLE problem consisting of two dependent stages: distributed detection and distributed estimation. The WSN performs distributed detection first and based on the global detection decision the distributed estimation stage is performed. The communication between the SNs and the FC occurs over a shared channel via a slotted Aloha MAC protocol to conserve BW.

In distributed detection, hard decision fusion is adopted, using the counting rule (CR), and sensor censoring in order to save power and BW. The effect of Rayleigh fading on distributed detection is also considered and accounted for by using distributed diversity combining techniques where the diversity combining is among the sensor nodes (SNs) in lieu of having the processing done at the FC. Two distributed techniques are proposed: the distributed maximum ratio combining (dMRC) and the distributed Equal Gain Combining (dEGC). Both techniques show superior detection performance when compared to conventional diversity combining procedures that take place at the FC.

In distributed estimation, the segmented distributed localization and estimation (SDLE) framework is proposed. The SDLE enables efficient power and BW processing. The SDLE hinges on the idea of introducing intermediate parameters that are estimated locally by the SNs and transmitted to the FC instead of the actual measurements. This concept decouples the main problem into a simpler set of local estimation problems solved at the SNs and a global estimation problem

solved at the FC. Two algorithms are proposed for solving the local problem: a nonlinear least squares (NLS) algorithm using the variable projection (VP) method and a simpler grid search (GS) method. Also, Four algorithms are proposed to solve the global problem: NLS, GS, hyperspherical intersection method (HSI), and robust hyperspherical intersection (RHSI) method. Thus, the SDLE can be solved through local and global algorithm combinations. Five combinations are tied: NLS² (NLS-NLS), NLS-HSI, NLS-RHSI, GS², and GS-NLS. It turns out that the last algorithm combination delivers the best localization and estimation performance. In fact, the target can be localized with less than one meter error.

The SNs send their local estimates to the FC over a shared channel using the slotted-Aloha MAC protocol, which suits WSNs since it requires only one channel. However, Aloha is known for its relatively high medium access or contention delay given the medium access probability is poorly chosen. This fact significantly hinders the time-critical operation of the system. Hence, multi-packet reception (MPR) is used with slotted Aloha protocol, in which several channels are used for contention. The contention delay is analyzed for slotted Aloha with and without MPR. More specifically, the mean and variance have been analytically computed and the contention delay distribution is approximated. Having theoretical expressions for the contention delay statistics enables optimizing both the medium access probability and the number of MPR channels in order to strike a trade-off between delay performance and complexity.

Contents

Abstract	vi
List of Figures	xiv
List of Acronyms and Symbols	xviii
1 Introduction	1
1.1 Motivation	1
1.2 Literature Overview	5
1.3 Thesis Proposal and Contributions	6
1.3.1 Proposal	6
1.3.2 Contributions	10
1.3.2.1 Distributed Detection	12
1.3.2.2 Distributed Estimation	12
1.3.2.3 Contention Delay	13
1.3.2.4 Publications	13
1.4 Preliminaries	14
1.4.1 Notation	14
1.4.2 Stochastic Geometry	15
1.4.3 Assumptions	16
1.5 Thesis Outline	17
2 Distributed Detection over Rayleigh Fading Multiple Access Channel	18
2.1 Introduction	18
2.1.1 Motivation	18
2.1.2 Related Work	20
2.1.3 Chapter Contributions	21
2.1.4 Chapter Outline	23
2.2 Problem Formulation	24
2.2.1 System Model	24

2.2.2	Distributed Detection	25
2.2.2.1	Local Detection	25
2.2.2.2	Global Detection	27
2.3	Stochastic Geometry Analysis of Detecting Sensor Nodes	28
2.4	Decision Fusion over Ideal Parallel Channels	33
2.4.1	Optimal Decision Fusion Rule	34
2.4.2	The Counting Rule	37
2.5	Decision Fusion over Shared MAC	38
2.5.1	Ideal Channel	39
2.5.2	Gaussian Channel	41
2.5.3	Rayleigh Fading Channel	43
2.6	Diversity Combining Techniques for MAC	45
2.6.1	Centralized Diversity Combining	45
2.6.2	Distributed Diversity Combining	46
2.6.2.1	Optimal Distributed Diversity Combining	47
2.6.2.2	Distributed Maximum Ratio Combining (dMRC)	49
2.6.2.3	Distributed Equal Gain Combining (dEGC)	51
2.7	Simulation Results and Discussions	52
2.7.1	Optimal Fusion Rule	52
2.7.2	Fusion Rules over Fading Shared Channel	53
2.8	Chapter Summary	63
3	Distributed Localization and Estimation of a Diffusive Target	66
3.1	Introduction	66
3.1.1	Motivation	66
3.1.2	Related Work	67
3.1.2.1	Static Parametric Model	67
3.1.2.2	Dynamical Parametric Model	69
3.1.2.3	Basis Representation Model	70
3.1.2.4	Convolutional Model	70
3.1.3	Chapter Proposal and Contributions	70
3.1.4	Chapter Outline	71
3.2	Problem Formulation	73
3.2.1	Diffusion Model	73
3.2.2	System Model	76
3.2.3	Distributed Estimation Problem	77
3.3	Classical Distributed Estimation	77
3.4	Cramer Rao Lower Bound	79
3.5	Segmented Distributed Localization and Estimation	82

3.5.1	Elements of SDLE	83
3.5.1.1	Reparameterization	84
3.5.1.2	Distributed Detection	85
3.5.1.3	Local and Global Processing	86
3.5.2	Local Estimation Techniques	88
3.5.2.1	Gradient-Based Search	88
3.5.2.2	Grid Search	90
3.5.3	Global Estimation Techniques	91
3.5.3.1	Gradient-based Search	91
3.5.3.2	Grid Search	92
3.5.3.3	Hyper-Spherical Intersection (HSI) Method	93
3.6	Simulation Results and Discussion	95
3.6.1	Simulation Setup	96
3.6.2	CRLB	99
3.6.3	MSE of SDLE Algorithms	99
3.7	Chapter Summary	103
4	Medium Access Delay for Slotted Aloha in Time-Critical Wireless Sensor Networks	108
4.1	Introduction	108
4.1.1	Motivation	108
4.1.2	Related Work	109
4.1.3	Chapter Proposal and Contributions	109
4.1.4	Chapter Outline	110
4.2	Problem Formulation	110
4.2.1	System Model	110
4.2.2	Contention Delay in DDLE	110
4.3	Contention Delay Analysis in Slotted Aloha	111
4.3.1	Exact Distribution	113
4.3.2	Negative Binomial Approximation	114
4.3.3	Medium Access Probability Impact	115
4.4	Contention Delay Analysis in MPR Slotted Aloha	116
4.5	Simulation Results and Discussion	118
4.5.1	Slotted Aloha	118
4.5.2	MPR Slotted Aloha	118
4.6	Chapter Summary	124
5	Overview, Conclusions, and Future Work	125
5.1	Conclusions	126
5.2	Future Work	128

5.2.1	Short term Work	128
5.2.1.1	Distributed Detection	128
5.2.1.2	Distributed Estimation	129
5.2.1.3	MAC Protocol	129
5.2.2	Long term Work	129
A	Proofs in Chapter 2	130
A.1	Proof of Theorem 2.1	130
A.2	Proof of Theorem 2.2	130
A.3	Proof of Lemma 2.1	132
A.4	Proof of Proposition 2.1	133
A.5	Proof of Proposition 2.2	134
A.6	Proof of Lemma 2.2	135
B	Derivation and Proofs in Chapter 3	136
B.1	Analytical Expressions for Derivatives in the Diffusion Case	136
B.2	Derivation of Fisher Information Matrix for an arbitrary Sensor Node	137
B.3	Proof of Theorem 3.1	138
B.4	Proof of Lemma 3.1	140
	References	144

List of Figures

1.1.1 WSN system diagram.	4
1.3.1 Centralized Detection and Estimation.	7
1.3.2 Distributed detection, localization, and estimation (DDLE) block diagram.	9
1.3.3 Distributed Detection, Localization and Estimation (DDLE) algorithm.	11
2.1.1 Distributed Detection, Localization and Estimation (DDLE) algorithm. In this chapter, section (I) will be discussed.	19
2.1.2 Parallel Access Channel WSN.	21
2.1.3 Multiple Access Channel WSN.	22
2.2.1 Distributed Detection in Wireless Sensor Networks System Diagram.	27
2.3.1 The averaged local probability of detection ($\mathbb{E}_A [P_d(\mathbf{x}, A)]$) plotted against the distance from the target ($\ \mathbf{x} - \mathbf{x}_t\ $).	32
2.7.1 The mean of the optimal fusion rule (OFR) ($\mathbb{E}[\Lambda_{\text{OFR}}]$) plotted against the local probability of false alarm, P_{fa} , under both null and alternative hypotheses, \mathcal{H}_0 and \mathcal{H}_1 respectively. The solid line represents simulated values and the dashed line is the theoretical values.	54
2.7.2 The variance of the optimal fusion rule (OFR) ($\mathbb{E}[\Lambda_{\text{OFR}}]$) plotted against the local probability of false alarm, P_{fa} , under both null and alternative hypotheses, \mathcal{H}_0 and \mathcal{H}_1 respectively. The solid line represents simulated values and the dashed line is the theoretical values.	55
2.7.3 ROC for optimal fusion rule (OFR) at $P_{fa} = 10^{-3}$. Solid line represents the Gamma distribution approximation and the dashed line represents the normal approximation.	56
2.7.4 ROC for optimal fusion rule (OFR) at $P_{fa} = 10^{-2}$. Solid line represents the Gamma distribution approximation and the dashed line represents the normal approximation.	57

2.7.5 Global probability of detection, P_D , against global probability of false alarm, P_{FA} . Solid line represents the theoretical values, and dashed line represents the simulation values. Network deployment density $\lambda = 5 \times 10^{-3}$, channel noise $\text{SNR}_c = -20\text{dB}$, Rayleigh channel power $\sigma_h^2 = 0\text{dB}$, and local detection threshold $\gamma = 235$ ($P_{fa} = 10^{-2}$).	59
2.7.6 Global probability of detection, P_D , against global probability of false alarm, P_{FA} . Solid line represents the theoretical values, and dashed line represents the simulation values. Network deployment density $\lambda = 5 \times 10^{-3}$, channel noise $\text{SNR}_c = -20\text{dB}$, Rayleigh channel power $\sigma_h^2 = 0\text{dB}$, and local detection threshold $\gamma = 235$	60
2.7.7 Probability of error, P_e , against local detection threshold, γ , at $\sigma_h^2 = 1$ and $\lambda = 1 \times 10^{-3}$	61
2.7.8 Probability of error, P_e , against transmission SNR_c , at $\sigma_h^2 = 1$ and $\lambda = 1 \times 10^{-3}$	62
2.7.9 Probability of error, P_e , against network deployment density, λ , with $\sigma_h^2 = 1$, and $\gamma = 235$	64
3.1.1 Distributed Detection, Localization, and Estimation (DDLE). In this chapter, section (II) is investigated.	72
3.5.1 Segmented distributed localization and estimation system diagram. $\hat{\vartheta}_i$ is an intermediate local estimate and $\hat{\theta}$ is the global estimate of the desired parameters.	83
3.5.2 Global Cost Function.	92
3.6.1 Wireless sensor network realization in sensing field \mathcal{F} , with target located at (1,-1) for $\lambda = 10^{-2}$	96
3.6.2 Concentration values plotted versus both time (t) and distance from target ($\ \mathbf{x}_t - \mathbf{x}\ $).	97
3.6.3 CRLB of θ given in (3.6.1) plotted against the sensing SNR (SNR_s) with $\gamma = 191$ corresponding to local probability of false alarm of 10^{-7}	100
3.6.4 CRLB of θ for values given in (3.6.1) plotted against γ with $\lambda = 3 \times 10^{-2}$ SN/ m^2	101
3.6.5 CRLB of θ given in (3.6.1) plotted against λ with $\gamma = 191$ corresponding to a local probability of false alarm of 10^{-7}	102
3.6.6 Impact of local detection threshold, γ , on the location of detecting SNs.	103
3.6.7 Local estimation RMSE of intermediate parameters, (a, b, n_r^i) , plotted against the local detection threshold γ , with $\lambda = 5 \times 10^{-2}$	104
3.6.8 Global estimation RMSE of θ plotted against the local detection threshold γ , with $\lambda = 5 \times 10^{-2}$	105
3.6.9 Average number of detecting sensor nodes plotted against the local detection threshold γ , with SN deployment density $\lambda = 5 \times 10^{-2}$	106
3.6.10 RMSE of estimators plotted against SN deployment density λ , with local detection threshold $\gamma = 172$ corresponding to local probability of false alarm of 10^{-5}	107

4.2.1 Distributed Detection, Localization and Estimation (DDLE) algorithm. In this chapter, section (III) is discussed.	112
4.5.1 Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 1 \times 10^{-2}$ and $\gamma = 172$ ($P_{fa} = 10^{-5}$).	119
4.5.2 Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 1 \times 10^{-2}$ and $\gamma = 118$ ($P_{fa} = 10^{-1}$).	120
4.5.3 Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 1 \times 10^{-3}$ and $\gamma = 172$ ($P_{fa} = 10^{-5}$).	121
4.5.4 Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 5 \times 10^{-3}$ and $\gamma = 172$ ($P_{fa} = 10^{-5}$).	122
4.5.5 Contention delay statistics plotted against the number of MPR channels for $\lambda = 1 \times 10^{-2}$ and $\gamma = 150$ ($P_{fa} = 10^{-3}$).	123

List of Acronyms and Symbols

Acronyms

BS	Base station
BW	Bandwidth
CDC	Centralized diversity combining
CDMA	Code division multiple access
CR	Counting rule
CRLB	Cramer Rao lower bound
CSI	Channel state information
CSMA	Carrier sense multiple access
DDLE	Distributed detection, localization, and estimation
dEGC	Distributed equal gain combining
dMRC	Distributed maximum ratio combining
ED	Energy detector
FC	Fusion center
FDMA	Frequency division multiple access
GLRT	Generalized likelihood ratio test
GN	Gauss-Newton method
GS	Grid search

HSI	Hyperspherical intersection
LM	Levenberg-Marquardt method
MAC	Multiple access channel. (In Chapter 4 it refers to Medium access control)
MLE	Maximum likelihood estimate
MPPP	Marked Poisson point process
NLS	Nonlinear least squares
OFR	Optimal fusion rule
OOK	On-off keying
PACs	Parallel access channels
PP	Point processes
PPP	Poisson point process
RHSI	Robust hyperspherical intersection
ROC	Receiver operation curve
RV	Random variable
SDLE	Segmented distributed localization and estimation
SNR	Signal to noise ratio
SNs	Sensor nodes
TBMA	Type based multiple access
TDMA	Time division multiple access
TOA	Time of arrival
VP	Variable projection
WSN	Wireless sensor network

Mathematical Symbols

$(\cdot)^T$	Transpose operator
*	Conjugate operator

θ	Vector of desired parameters
ϑ_i	Intermediate parameter of the i th SN
\mathbf{X}_i	Sensor node location
\mathbf{X}_t	Random target location
$\delta(\cdot)$	Kronecker delta function
κ	Diffusivity
Λ	Global detection test statistic
$\ \cdot\ $	Euclidean norm
$\ \cdot\ _F$	Frobenius norm
$\mathbb{E}[\cdot]$	Expectation operator
$\mathbb{P}(\cdot)$	Probability
\mathbb{R}	Set of real numbers
\mathbb{R}^2	xy-plane
\mathbf{x}_t	Target's location realization
\mathbf{x}	Space coordinate
\mathcal{F}	Sensing field
\mathcal{H}_0	Target absent hypothesis
\mathcal{H}_1	Target present hypothesis
μ	Release rate
∇^2	Spatial second derivative
Φ	Point process
$\Re\{\cdot\}$	Real operator
σ_c^2	Communication channel noise variance
σ_h^2	Fading channel variance
σ_s^2	Sensing measurements noise variance
τ	Target's entry time

$\text{var}[\cdot]$	Variance operator
$\text{erfc}(\cdot)$	Complementary error function
$\text{tr}\{\cdot\}$	Matrix trace
Υ	Global detection threshold
$\hat{\theta}$	Estimate of desired parameters vector
Ξ	Test statistic
I_{GD}	Global detection decision
I_i	Local detection decision
M	Number of SNs
N	Number of samples
P_D	Global probability of detection
P_d	Local probability of detection
P_{FA}	Global probability of false alarm
P_{fa}	Local probability of false alarm
t	Time
T_s	Sampling period

Chapter 1

Introduction

1.1 Motivation

Wireless sensor networks (WSNs) are considered one of the most important emerging technologies in the 21st century [1] and one of ten technologies that will change the world [2]. WSNs are envisioned to be composed of cheap, smart devices with multiple on-board sensors that are deployed in large numbers and wirelessly networked together. This endows the WSN with scalability, versatility, and flexibility. The sensor node (SN) simply consists of a transceiver, sensing device(s), power storage (batteries usually), and a processor [3, 4]. This structure enables the WSN to be densely dispersed over a large geographically area in large numbers, thus providing deployment scalability and robustness to node failures. But limited on-board power restricts the system lifetime and communication BW. Cheap hardware also limits the computation ability. Furthermore, relying on wireless channels introduces reliability issues. Nevertheless, WSNs are used for a variety of applications [5] such as:

1. **Environmental applications:** Natural habitat monitoring of the Great Duck island using WSNs was reported in [6]. Under project *SensorScope* [7], a WSN was used for atmospheric monitoring of a mountainous area between Switzerland and Italy. In [8], WSNs were also used in monitoring, automating, and controlling an agricultural greenhouse. Volcanic eruptions were monitored using a WSN in central Ecuador [9].
2. **Industrial applications:** A WSN was instrumented in preventive equipment maintenance in [10] for a semiconductor plant and shipboard engine room.

1.1. MOTIVATION

Bridges and tunnels were monitored using a large scale WSN deployment [11].

3. **Civilian alarming applications:** SensorScope was also used in a building monitoring system [12]. A landslide detection system based on a WSN was discussed in [13]. A real-time fire rescue application was presented in [14]. Border monitoring employing WSNs was investigated in [15] using the *BorderSense* system.
4. **Military applications:** One of the earliest implementation of a WSN in a military application was in the *SensIT* project [16] for target detection and tracking. Another military surveillance system using a WSN was carried out in [17] to detect and track enemy vehicles using 70 MICA2 nodes [18]. Similarly, a WSN system for multi-target tracking was presented in [19] using 557 Triomotes [20]. Even larger WSN deployment of over 1000 SNs was introduced in the *ExScal* project [21], which was derived from an earlier field demonstration [22]. Sniper detection based WSN was proposed in [23, 24], which was based on the work in [25]. The *VigilNet* project [?] also took on the task of implementing real-time tracking using WSNs, where the detection and classification algorithms were based on [26].

The above applications can be sorted into two general types: monitoring and surveillance. In the former, WSNs are used to gather information about the surrounding environment and send it to a data sink. There, the main purpose is to maintain a satisfactory packet delivery rate at the sink while preserving a sufficient system lifetime. Thus the WSN is designed as a distributed data acquisition system with appropriate medium access control (MAC) [27] and routing protocols [28] in order to achieve the desired packet delivery rate with limited power and BW budget. Thus the emphasis in monitoring is on the communication side of the network. Whereas in surveillance WSN, the main objective is to detect the presence of any targets, track their location, and classify their type (i.e., person, vehicle, or animal) within a specific time constraint (e.g., before the target leaves the designated area of interest) [29, 30]. Hence, both signal processing and communication should be considered in the system design, which is more challenging because the signal processing algorithms must be adapted to the restricted resources available. Note however, that target tracking requires repeated localization and the classification requires estimation of the target's attributes which are checked against a predetermined profile. Therefore, we will focus on the localization and estimation operations in surveillance applications.

An illustrative example of the surveillance application would be using WSNs to guard and monitor gas pipelines. For example the gas pipeline carrying the Egyptian

1.1. MOTIVATION

gas to Jordan through the Sinai desert. However, this pipeline has been bombed more than fifteen times since 2011 [31]. It is virtually impossible to efficiently guard this line using fences, manned patrols, or even surveillance cameras. For instance, constructing fences around the pipeline is highly expensive since the pipeline extends over very long distances. Whereas manned patrols cannot provide around the clock security. Surveillance cameras on the other hand, require infrastructure to provide power and communication for the system, which is significantly challenging in harsh environments such as deserts. In such situation however, WSNs provide an efficient reasonable solution, since they can be deployed over vast geographical areas to perform intruder detection and localization tasks via seamlessly forming ad-hoc networks, and sustaining themselves via on-board batteries or energy harvesting. However, it is still required to have a series of base stations located far away from the network to act as fusion centers to cover the whole area of interest.

To appreciate the existing challenges in surveillance WSN, we take a deeper look into a local area covered by a single FC as shown in Fig. 1.1.1. When a target enters the sensing field, it generates a signature in one or more modalities (e.g., thermal, acoustic, magnetic, or a chemical trace). The SNs collect measurements from this physical phenomenon, process them, and then transmit this data wirelessly to the fusion center (FC) for target detection, localization, and estimation. Conventional centralized detection and estimation algorithms, in which all the data is processed at a single point in the network, do not conform with the WSN's constraints and the real-time surveillance requirement. From a communication point of view, sending all the data to the FC leads to excessive use of energy and bandwidth, both of which are scarce resources in WSNs. Moreover, such a set up will significantly burden the MAC protocol especially for large networks. In the case of deterministic MAC protocols (such as TDMA, FDMA, and CDMA) it would result in high complexity due to the large time and frequency synchronization requirements. Hence using contention based MAC protocols (such as Aloha and CSMA) might be more resource efficient. However such protocols are known to introduce high delays, which could compromise the time critical operation of the WSN. Although other factors contribute to the overall delay like queuing and routing, we concentrate on the MAC delay here because it generally affects other elements in the network. For example, in queuing, the service rate of the queue is actually the successful medium access rate. The routing delay is actually the aggregation of the queuing and MAC delay in the case of multiple-hop network. From a signal processing point view however, the centralized approach requires high computational capacity. On the one hand, the wireless channels between the SNs and the FC need to be equalized. On the other hand, detection and estimation algorithms using a large amount of data are

1.1. MOTIVATION

rather complicated.

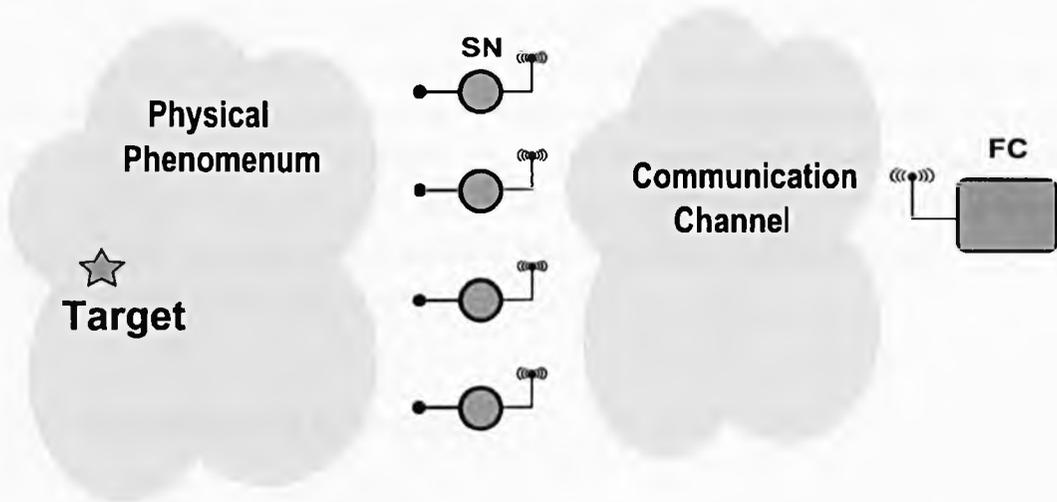


Figure 1.1.1: WSN system diagram.

Distributed processing is posed as an efficient alternative in which the signal processing is distributed among the SN and the FC. There are two kinds, generally, of distributed processing: FC-based and Ad-hoc. In the former, the processing is divided between the SNs and the FC. While in the latter, there is no central point in the network, i.e., all the processing is done locally at the SNs. In both cases, distributed processing is more resource efficient because local processing reduces the amount of data disseminated in the network, which in turn leads to less power and BW expenditure.

In this thesis, we consider the problem of distributed detection, localization, and estimation (DDLE) of a stationary target in a FC-based WSN subjected to time-critical operation and restricted power and BW resources. We also consider transmission over a shared channel suffering from Rayleigh fading. To solve this problem, we propose a novel efficient modular DDLE algorithm in which the main detection and estimation problems are decoupled into simpler and cheaper local and global problems. The distributed detection and estimation are divided into separate tasks that are performed sequentially, i.e., the WSN performs distributed detection first and based on the global detection decision the distributed estimation is performed. In distributed detection, we adopt hard decision fusion and sensor censoring [32] in order to save power and BW. Moreover, using hard decisions facilitates the use of a shared channel for communication¹. The effect of Rayleigh fading

¹This statement is thoroughly investigated in Chapter 2.

1.2. LITERATURE OVERVIEW

on distributed detection is also considered and accounted for by using novel distributed diversity combining techniques. In distributed estimation, we propose the segmented distributed localization and estimation (SDLE) framework that enables efficient power and BW processing. The SDLE relies on the idea of introducing intermediate parameters that are estimated locally by the SNs and transmitted to the FC instead of the actual measurements. The FC then uses the intermediate parameters to estimate the original target's parameters. In order to transmit the intermediate estimates efficiently and with acceptable delay, we adopt a multiple-packet reception (MPR) slotted Aloha MAC protocol.

1.2 Literature Overview

Several interesting attempts to address this problem has been devised in the literature. However, most of the work addresses the detection and the estimation problems separately. The literature is abundant with theoretical work on distributed detection [33, 34, 35, 36, 37, 38, 39, 40, 41] and distributed estimation [42, 43, 44, 45, 46, 47, 48, 49, 50, 51]. However, the previous treatments tend to have used simplified assumptions about the target and the channel, such as a constant target signal and an ideal channel. Therefore, we prefer to look at practical surveillance WSN systems in order to understand the actual problems and constraints facing such real-world implementations².

In the VigilNet project [?], the real-time tracking problem was approached by dividing the system into several phases and minimizing the corresponding time of each phase to satisfy the overall real-time requirement. VigilNet adopts a hierarchical paradigm in which the the SNs group together and report to group leaders that also report back to a base station (BS). The classification algorithm there relies on the SNs sending a *confidence vector* to the group leader, which is statically chosen. This confidence vector has confidence levels of detections from different sensing modalities such as acoustic, motion (using pyroelectric infra-red (PIR) sensors), and magnetic. The sensors samples are filtered and then fed to energy detectors (EDs) to compute the confidence vectors. The group leader then aggregates the confidence vectors from the SNs into a group-level confidence report that is in turn send to the BS where the ultimate classification takes place.

The ExScal project [21] also adopts a hierarchical topology. However, the network is divided into Tier 1 consisting of XSM (Extreme Scale Mote) nodes responsible for sensing and detection, Tier 2 XSS (Extreme Scale Stargate) nodes that

²Deployment problems, on the other had, were thoroughly discussed in [52, 53, 54].

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

form the communication backbone, and Tier 3 node, which is the master operator or a BS. Tier 1 nodes employ simple signal processing techniques similar to the VigilNet case. The classification algorithm though, depends on the concept of *field of influence* that is actually the area around the target having detecting SNs. Different targets have a different field of influence, hence different number of detecting SNs. Therefore, the BS can classify a target based on the detection reports received from Tier 1 via Tier 2 nodes. Localization is performed via the centroid method however.

In contrast to VigilNet and ExScal, which use low end signal processing, the Urban sniper localization system [23, 24] employs more powerful signal processing like time-frequency analysis to extract the time of arrival (TOA) at each SNs. The TOAs then are sent to the BS, via multihop transmission, to be fused together by maximizing a *consistency function* appropriately defined to find the final sniper position. It is noteworthy though, that sophisticated hardware such as a digital signal processor (DSP) and field programmable gate array (FPGA) were used to implement the TOA processing in the SNs.

Two remarks can be drawn from the above projects. First, if no analytical signal model is specified, simple signal processing methods are used (e.g., energy detector or FFT) for local processing in the SNs. The loss of performance though, is compensated by the large numbers of SNs involved. Second, when the signal model is specified, more complex signal processing is needed, which leads to more complicated and expensive hardware used for local processing. This limits the use of SNs in large numbers to improve performance.

1.3 Thesis Proposal and Contributions

1.3.1 Proposal

In the light of the above conclusions about practical WSN implementations, we desire to have a sophisticated signal model that describes the *spatio-temporal* variation of the measured signal, while in the mean time keep the local (and global) signal processing as simple as possible in order to leverage the power of (SN) numbers in the WSN.

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

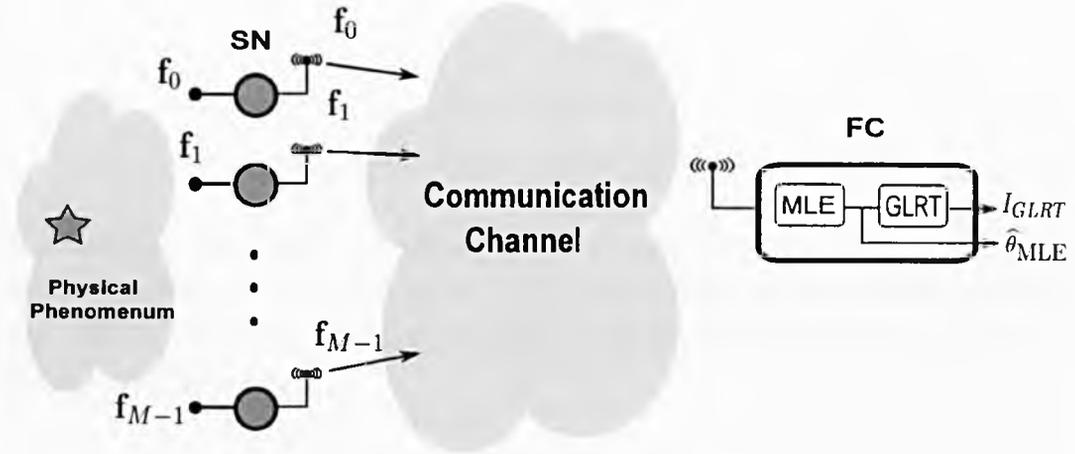


Figure 1.3.1: Centralized Detection and Estimation.

Suppose there is a WSN with M SNs deployed in sensing field \mathcal{F} for surveillance as shown in Fig. 1.3.1. A target enters the sensing field at an arbitrary time and location. This target leaves a signature signal $s(\mathbf{x}, t; \boldsymbol{\theta})$, which varies with space (\mathbf{x}) and time (t). The signature is characterized by a set of parameters compiled in vector $\boldsymbol{\theta}$ that includes the target's location and entry time. The i th SN located at \mathbf{x}_i collects N noisy measurements

$$f_i[n] = s[\mathbf{x}_i, n; \boldsymbol{\theta}] + w_i[n] \quad (1.3.1)$$

where $w_i[n]$ is the n th noise sample at the i th SN.

In contrast to conventional space-invariant models that are popular in the literature, such as the constant-plus-noise ($s[n; \boldsymbol{\theta}] = \theta$) and the regression model ($s[n; \boldsymbol{\theta}] = \mathbf{H}\boldsymbol{\theta}$), we adopt spatio-temporal signal models to better reflect practical scenarios. In fact, we discuss two signature models for targets in this thesis. The first is the power law, which is used in the distributed detection discussion since it emphasizes the spatial variation of the signal. The second model is the diffusion law [55], which is used in the distributed estimation discussion. The diffusion signal varies with space and time and is classified as a non-wave field [56], which is more general than the conventional sensor array models. However, the analysis in both distributed detection and estimation is fairly general and can be applied to different models as well.

Using $f_i[n]$ s for all $i = 0, \dots, M-1$ and $n = 0, \dots, N-1$, it is desired to detect the presence of the target and estimate its parameters efficiently. In other words, we wish to test the following hypotheses:

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

$$\mathcal{H}_1 : \text{Target is present} \quad (1.3.2)$$

$$\mathcal{H}_0 : \text{Target is NOT present} \quad (1.3.3)$$

and then find the target's estimate ($\hat{\theta}$) if the target is actually present. The conventional approach, as shown in Fig. 1.3.1, is to use the generalized likelihood ratio test (GLRT) [57] that requires finding the maximum likelihood estimate (MLE) of θ , say $\hat{\theta}_{\text{MLE}}$, by solving

$$\hat{\theta}_{\text{MLE}} = \arg \max_{\theta} l(\mathbf{F}, \theta; \mathcal{H}_1) \quad (1.3.4)$$

where $\mathbf{F} = (\mathbf{f}_0, \mathbf{f}_1, \dots, \mathbf{f}_{M-1})$ is an $N \times M$ matrix, $\mathbf{f}_i = (f_i[0], f_i[1], \dots, f_i[N-1])^T$ is $N \times 1$ measurement vector from the i th SN, $(\cdot)^T$ is the transpose operator, and $l(\mathbf{F}, \theta; \mathcal{H}_1)$ is the likelihood function under the \mathcal{H}_1 hypothesis. The $\hat{\theta}_{\text{MLE}}$ is then substituted into the likelihood ratio, i.e.,

$$\Lambda_{\text{GLTR}} = \frac{l(\mathbf{F}, \hat{\theta}_{\text{MLE}}; \mathcal{H}_1)}{l(\mathbf{F}; \mathcal{H}_0)}. \quad (1.3.5)$$

The global detection decision is found by comparing the above with a specific threshold, say Υ , i.e.,

$$I_{\text{GLRT}} = \begin{cases} 1, & \Lambda_{\text{GLTR}} \geq \Upsilon \\ 0, & \Lambda_{\text{GLTR}} < \Upsilon. \end{cases} \quad (1.3.6)$$

Obviously, the GLRT leads to excessive use of power and BW because (\mathbf{F}) needs to be sent to the FC that also causes large MAC delays. Furthermore, solving (1.3.4) is computationally demanding. However, such conditions do not suit WSNs operation.

In this thesis, we efficiently solve the DDLE problem by dividing the problem into distributed detection and distributed estimation that are dependently executed and then analytically evaluate the performance. The proposed DDLE system is shown in Fig. 1.3.2. In contrast to the conventional approach shown in Fig. 1.3.1, the detection is performed before the estimation in both local and global processing. This feature enables the WSN to save energy and BW. However, this efficiency comes with price of deviating from the optimal GLRT approach.

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

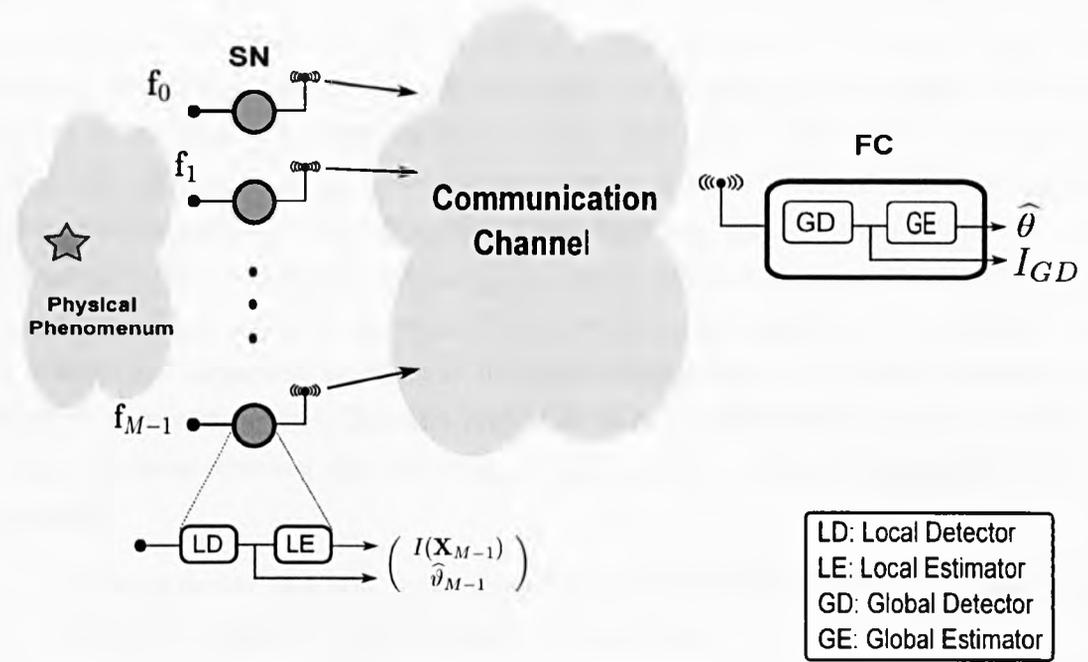


Figure 1.3.2: Distributed detection, localization, and estimation (DDLE) block diagram.

To understand our proposed algorithm, we turn to Fig. 1.3.3 that shows the proposed DDLE procedure. The algorithm starts with network initialization, the FC can send a known training sequence to the SNs to be used to perform the necessary channel estimation and synchronization [58], in addition to node localization [59]. Then distribution detection is performed by having the SNs sample the sensing medium and collect measurements periodically. Those measurements are used to perform local detection at the SN to come up with a local decision (I_i) at the i th SN. If the local decisions are positive ($I_i = 1$), they are then sent to the FC over a shared channel where they are fused³ to compute the global decision, Λ , according to the counting rule [38], which is simple to implement and analyze. This approach has several advantages:

- It is power and BW efficient since it only sends the positive decision (a binary one) to the FC instead of the measurements.
- It circumvents the need to estimate θ for detection purposes (in contrast to the GLRT), hence simplifying the local detector.
- Having local and global detection significantly reduces the false alarms and consequently saves the power and BW used in unnecessary estimation.

³When different signals are sent over a shared channel to the same receiver, they are naturally mixed (fused) together. This fact was investigated and used in [60].

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

If $I_i = 0$, the SNs do not send their data to the FC. Similarly, if the global decision is negative, the FC signals the SNs to return to collecting data. However, if the global decision is positive ($I_{GD} = 1$), which asserts the presence of the target, the WSN moves on to distributed estimation. In this stage, local intermediate parameters (ϑ'_i 's) are estimated at the SNs and then send to the FC where they are used to compute the desired estimate of θ . Upon finishing the estimation, the FC also signals the SNs to return to collecting measurements and hence restarting the whole algorithm. How often is the algorithm carried out is application dependent, e.g., if vehicle are expected to intrude into the sensing field, then the corresponding frequency should be such that the vehicle does not significantly change its location. This constraint enables the tracking of such targets. This paradigm also has its benefits:

- It saves power and BW by sending a small number of intermediate parameters to the FC instead of the collected measurements.
- It decouples the high dimensional MLE problem (1.3.4) into simpler local and global problems, which are solved at the SNs and FC respectively.

The local estimates (ϑ'_i 's) are transmitted to the FC over a shared channel as well, in which the MPR slotted Aloha MAC is adopted to ensure that they are received, in addition to having network scalability, transceiver simplicity, and resource efficiency⁴. Moreover, since the local parameters occupy a smaller bandwidth, powerful coding can be used to improve error correction at the FC. It is assumed however, that the local estimates are perfectly received at the FC.

The performance of the proposed DDLE algorithm is analyzed using stochastic geometry tools [61]. For distributed detection, the receiver operation curve (ROC) and the probability of error are characterized. Whereas in the distributed estimation case, the Cramer Rao lower bound (CRLB) is analytically derived and evaluated. The contention delay is also statistically analyzed.

1.3.2 Contributions

In addition to the previously proposed algorithm, we make the following contributions in distributed detection, distributed estimation, and contention delay analy-

⁴Note that in contrast to sending the local detection decisions (I_i 's), which are always the same signal, sending the local estimates requires arbitration of the communication channel for a single transmitting SN. Hence, a MAC protocol is indeed needed.

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

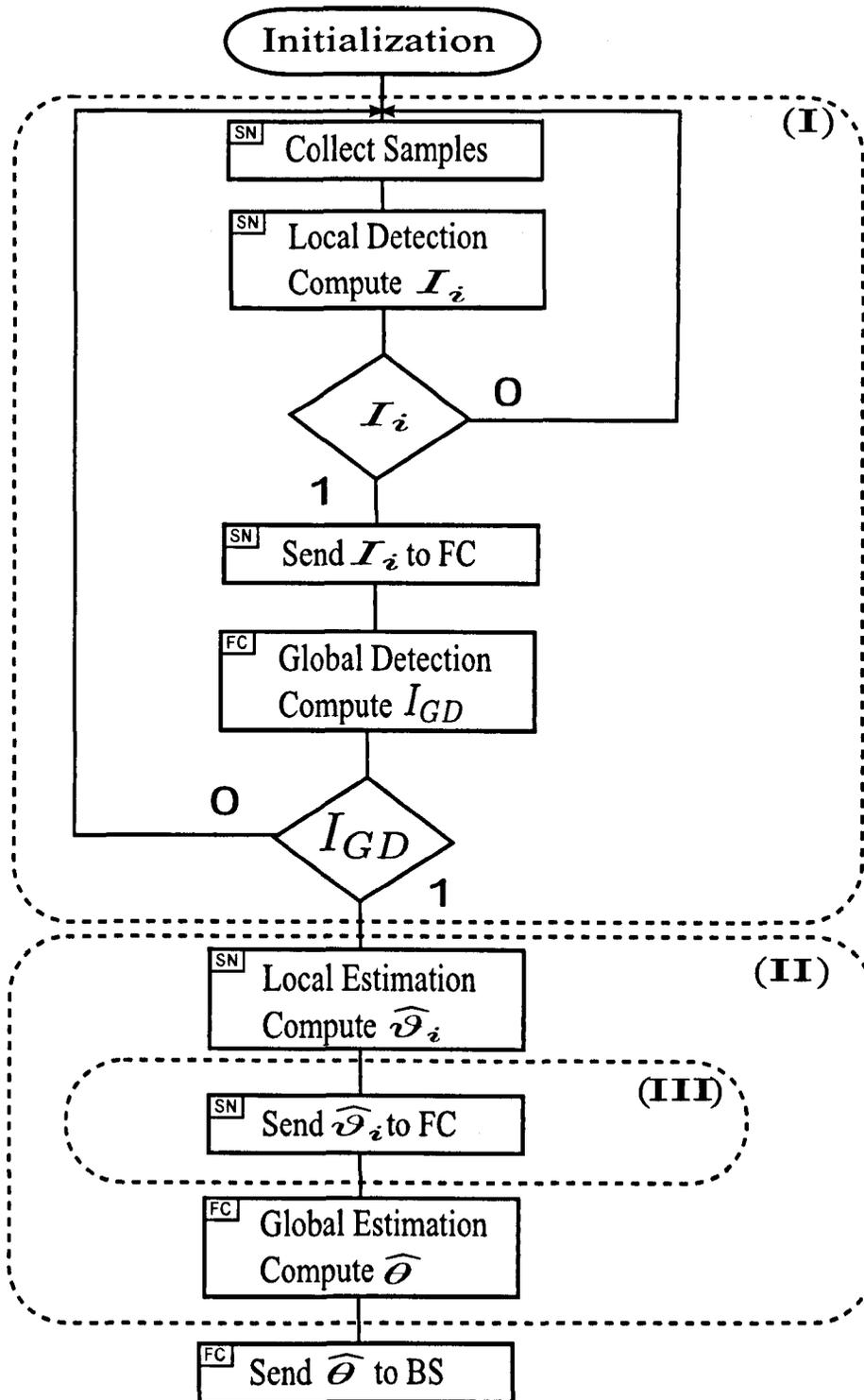


Figure 1.3.3: Distributed Detection, Localization and Estimation (DDLE) algorithm.

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

1.3.2.1 Distributed Detection

1. Using a stochastic geometry framework, the spatial distribution of the detection SNs has been characterized in theorems 2.1 and 2.2.
2. Using the above theorems we have found the theoretical performance of the optimal Chair-Varshney rule [62] and the counting rule [37, 38].
3. When distributed detection occurs over a shared Rayleigh fading channel, we have proposed and statistically analyzed the performance of the following distributed diversity combining algorithms:
 - a) Distributed Maximum Ratio Combining (dMRC).
 - b) Distributed Equal Gain Combining (dEGC).

that are shown to have better performance when compared to other rules in the literature.

1.3.2.2 Distributed Estimation

1. A SDLE framework has been introduced by using intermediate parameters. The original estimation problem was decoupled into a simpler more resource efficient set of local estimation problems and a global estimation problem, which both are nonlinear least squares (NLS). The following algorithms have been proposed to solve the local estimation problem:
 - a) Variable projection (VP) method.
 - b) Grid search (GS), which can be implemented as a simple correlator.

The following algorithms have been suggested to solve the global estimation problem by:

- a) Global VP,
- b) Global GS,
- c) Hyperspherical intersection (HSI) method,
- d) Robust hyperspherical intersection (RIISI) method.

We try out five combinations, NLS², NLS-HSI, NLS-RHSI, GS², and GS-NLS. It turns out that the last algorithm combination delivers the best localization and estimation performance.

1.3. THESIS PROPOSAL AND CONTRIBUTIONS

2. Using stochastic geometry, the CRLB was analytically computed. This expression is general and can be used for different signal models.

1.3.2.3 Contention Delay

We have defined the contention delay in DDLE WSN and statistically characterized the contention delay in MPR slotted-Aloha used for distributed detection. More specifically:

1. The mean and variance have been analytically computed for a single channel slotted-Aloha, via stochastic geometry tools and results from Chapter 2.
2. Using the above results, the contention delay distribution for a single channel slotted-Aloha has been approximated by a negative binomial distribution via moment matching.
3. The contention delay in MPR slotted-Aloha was also characterized by using order statistics and results from the single channel case.

Those theoretical results have then been used to find the optimal medium access probability and number of MPR channels, which minimizes the delay statistics.

1.3.2.4 Publications

Portions of this thesis have been presented at the:

- 2010 IEEE Seventh International Symposium on Wireless Communication Systems (ISWCS) [63].
- 2011 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP) [64].
- 2011 IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC) [65].
- 2012 IEEE wireless advanced (WiAd) [66].

Furthermore, we have submitted to the:

- IEEE Transaction on Signal Processing [67].
- IEEE Transaction on Vehicular Technology [68].

1.4 Preliminaries

1.4.1 Notation

We use upper case to denote a random variable (RV) and the lower case for its realization, except for the indexes i, j, k and n , and number of samples N . Bold variables represent vectors, e.g., \mathbf{x} is a deterministic vector whereas \mathbf{X} stands for a random vector. \mathbb{R}^d stands for a d -dimensional space. The vector transpose is denoted by the superscript $(\cdot)^T$. The vector norm used here is the Euclidean norm $\|\cdot\|$ unless stated otherwise. The probability distribution for a RV A is denoted by $p_A(a)$. Whereas $\mathbb{P}(B)$ is the probability of event B . $\mathcal{N}(\mu, \sigma^2)$ is a Gaussian distribution with mean μ and variance σ^2 . $\text{Poi}(\lambda)$ refers to the Poisson distribution with mean λ , i.e.,

$$\text{Poi}(\lambda) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Similarly, $\text{Gamma}(\cdot, \cdot)$, $\text{Geo}(\cdot)$ and $\text{NB}(\cdot, \cdot)$ refer to the Gamma, geometric, and negative binomial distributions respectively, are defined as:

$$\text{Gamma}(\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad (1.4.1)$$

$$\text{Geo}(q) = (1 - q)^k q \quad (1.4.2)$$

$$\text{NB}(r, q) = \binom{k+r-1}{k} (1-q)^r q^k \quad (1.4.3)$$

where $\Gamma(\cdot)$ is the Gamma function.

The expectation and variance under the j th hypothesis are $\mathbb{E}_j[\cdot]$ and $\text{var}_j[\cdot]$ respectively. However, a different subscript might be used to explicitly refer to the RV over which the expectation is taken, e.g. $\mathbb{E}_A[\cdot]$, as will be clear from the context. Another use of upper and lower case, this time in the subscripts, is to indicate being local or global. For example, P_d and P_{fa} are the local detection and false alarm probabilities respectively, while P_D and P_{FA} are the global detection and false alarm probabilities respectively.

1.4. PRELIMINARIES

1.4.2 Stochastic Geometry

Stochastic geometry is a mathematical tool that is becoming popular in modeling wireless networks [69, 70]. Here we provide a brief introduction to stochastic geometry. A more detailed discussion can be found in [61] and [71].

Stochastic geometry is mainly concerned with addressing point processes (PPs) and associated geometrical properties. A PP Φ is the set of points $\{\mathbf{X}_i\}$ random located in some space \mathbb{R}^d of dimension $d \geq 1$, where \mathbf{X}_i is a RV that represents the location of the i th point. We are interested in the \mathbb{R}^2 plane here, so we restrict our attention to $d = 2$. Φ is called simple if no two points occupy the same location. It is stationary if its statistical properties are unchanged after shifting. Similarly, Φ is called isotropic if its properties are unchanged after rotation. An interesting PP class is the stationary Poisson point process (PPP) that is characterized by two main properties (informally stated):

1. For a closed subset $\mathcal{F} \subset \mathbb{R}^2$, the number of points in it follows a Poisson distribution.
2. The points in a PPP are independently scattered. In other words the number of points in disjoint sets are independent.

We restrict the discussion now to PPP, however it also applies to a general PP. A PPP can be *marked* as well, by attaching a certain attribute or mark to each point in the PPP forming a marked PPP (MPPP). For example, $\Phi_m = \{(\mathbf{X}_i, M_i)\}$, where M_i is the mark of the i th point. Those marks usually represent some relevant quantity, e.g. the channel coefficient of a communication link between a SN and the base station. Marks are usually RVs drawn from a certain distribution $p_M(m)$.

As with any stochastic process, we are interested in the moments of Φ . The mean of the PP Φ is defined for a certain subset⁵ \mathcal{F} as $\mathbb{E}[\Phi(\mathcal{F})]$, which is the number of points in \mathcal{F} . For an infinitesimally small \mathcal{F} , the mean is called the *intensity* of Φ , which we denote as λ . The intensity can be constant, which leads to a *homogenous* PPP or space varying (i.e., $\lambda(\mathbf{x})$ for an arbitrary location $\mathbf{x} \in \mathcal{F}$) leading to an inhomogeneous PPP. Note that a homogenous PPP implies that the points are uniformly distributed in \mathbb{R}^2 in contrast to the inhomogeneous PPP. The second central moment of Φ , on the other hand, is defined as $\mathbb{E}[\Phi^2(\mathcal{F})] - \mathbb{E}^2[\Phi(\mathcal{F})]$.

Let us define a positive function on the point process $f(\mathbf{x})$ for an arbitrary location $\mathbf{x} \in \Phi$. Given that Φ is a stationary MPPP, then using Campbell's theorem

⁵ \mathcal{F} is an area in \mathbb{R}^2 and a volume in \mathbb{R}^3 . For the general \mathbb{R}^d space, \mathcal{F} is a Borel set, which is a generalization of volume in higher dimensions.

1.4. PRELIMINARIES

[70] we can get the mean and variance as

$$\mathbb{E} \left[\sum_{\mathbf{X}_i \in \Phi} f(\mathbf{X}_i) \right] = \int_{\mathbb{R}} \int_{\mathbb{R}^2} f(\mathbf{x}) \lambda(\mathbf{x}) p_M(m) d\mathbf{x} dm \quad (1.4.4)$$

$$\text{var} \left(\sum_{\mathbf{X}_i \in \Phi_m} f(\mathbf{X}_i) \right) = \int_{\mathbb{R}} \int_{\mathbb{R}^2} f^2(\mathbf{x}) \lambda(\mathbf{x}) p_M(m) d\mathbf{x} dm. \quad (1.4.5)$$

Campbell's theorem can also be used to find the mean and variance of a stationary PPP Φ if we set $p_M(m)$ to unity for all m and $f(\mathbf{X}) = \mathbf{1}(\mathbf{X})$, which is the indicator function of the PP locations, i.e., it takes unity value when $\mathbf{X} \in \Phi$ and zero otherwise.

Finally, we discuss the *thinning* operation on Φ . Thinning is actually the removing of points from Φ according to a certain regime. It can be deterministic or random. We are interested in the latter case here. In a random thinning, the point \mathbf{X}_i is retained with probability $P(\mathbf{X}_i)$. If the thinning operation depends on the location of the point, then it is called *dependent thinning*, otherwise it is known as *independent thinning*.

1.4.3 Assumptions

Throughout the thesis we assume that the following assumptions hold:

- The SNs locations are known to the FC.
- The SNs are synchronized to the FC.
- The channel states are perfectly estimated at the SNs.
- Given the above assumptions, we also assume that the local estimates sent to the FC are received perfectly.

Those are plausible assumptions since the first can be satisfied in the network initialization phase. The last two conditions can be fulfilled by using a periodic training sequence broadcast by the FC to all SNs.

1.5 Thesis Outline

The rest of this thesis is dedicated to discussing the DDLE algorithm in Fig. 1.3.3, which is divided into the following operations:

- (I) Distributed detection.
- (II) Distributed estimation.
- (III) Contention-based transmission.

Chapter 2 discusses the operation (I) in DDLE that involves distributed detection over a shared Rayleigh fading channel. The optimal hard decision fusion is analyzed first followed by the suboptimal counting rule. Both previous cases are under ideal parallel channels. Fusion rules over imperfect channels are then discussed and analyzed as well. This sets up the stage for the proposing and analyzing diversity combining techniques, dMRC and dEGC. These improve global detection, when compared to conventional diversity combining techniques.

Chapter 3 handles operation (II) that introduces the distributed localization and estimation of a diffusive target⁶. First, physical and statistical models for a diffusive target are presented. Then, the optimal distributed estimation is presented in the context of a random network modeled by a PPP. The CRLB is analytically found for the corresponding estimation problem using stochastic geometry. After that, the SDLE is suggested to efficiently implement distributed estimation, and five SDLE variants are presented based on the NLS, GS, HSI, and RHSI algorithms.

In Chapter 4 we investigate operation (III) of the DDLE that involves the medium access delay incurred at the FC to receive all the local estimates from the active SNs. The mean and variance of the contention delay are derived analytically using stochastic geometry tools. Then the statistical distribution is approximated by a negative binomial distribution, which is the discrete analog of the Gamma distribution. Note however, that even though the contention-based transmission is a part of the distributed estimation procedure, it is discussed before distribution estimation because it directly relies on the results provided by Chapter 2.

Finally, the thesis is concluded with Chapter 5 which contains conclusions and future work for the short and long terms.

⁶i.e., the target whose signature follows the diffusion law.

Chapter 2

Distributed Detection over Rayleigh Fading Multiple Access Channel

2.1 Introduction

2.1.1 Motivation

WSNs are naturally suited for the task of target detection. On the one hand, the SNs are distributed geographically to monitor vast areas for any intrusion by adopting wireless communication. On the other hand, having multiple SNs sending information to the FC provides *detection diversity*, which significantly improves the detection performance. However, transmitting information over a wireless channel raises the issue of channel imperfections such as fading. Furthermore, the low cost SNs put stringent constraints on the power and bandwidth use. An efficient approach, to cope with resource scarcity, is having the SNs perform local detection first, and then send their decisions to the FC where those decisions are fused into the global decision about the target's presence/absence. Conventionally, the local decisions are assumed to be sent over orthogonal channels. This assumption however, burdens the FC with the task of setting up an independent channel for each SN in the network, which can cause scalability issues if the network is large. Furthermore, this chore gets more complicated when the number of SNs is unknown, which is usually the case in WSNs.

In this chapter, we will address the problem of distributed detection in WSNs, which is the highlighted section (I) in Fig. 2.1.1 in the DDLE algorithm. The main

2.1. INTRODUCTION

issues there are to perform energy and BW efficient decision fusion and overcome the effect of the fading channel on the transmission of the SNs decisions.

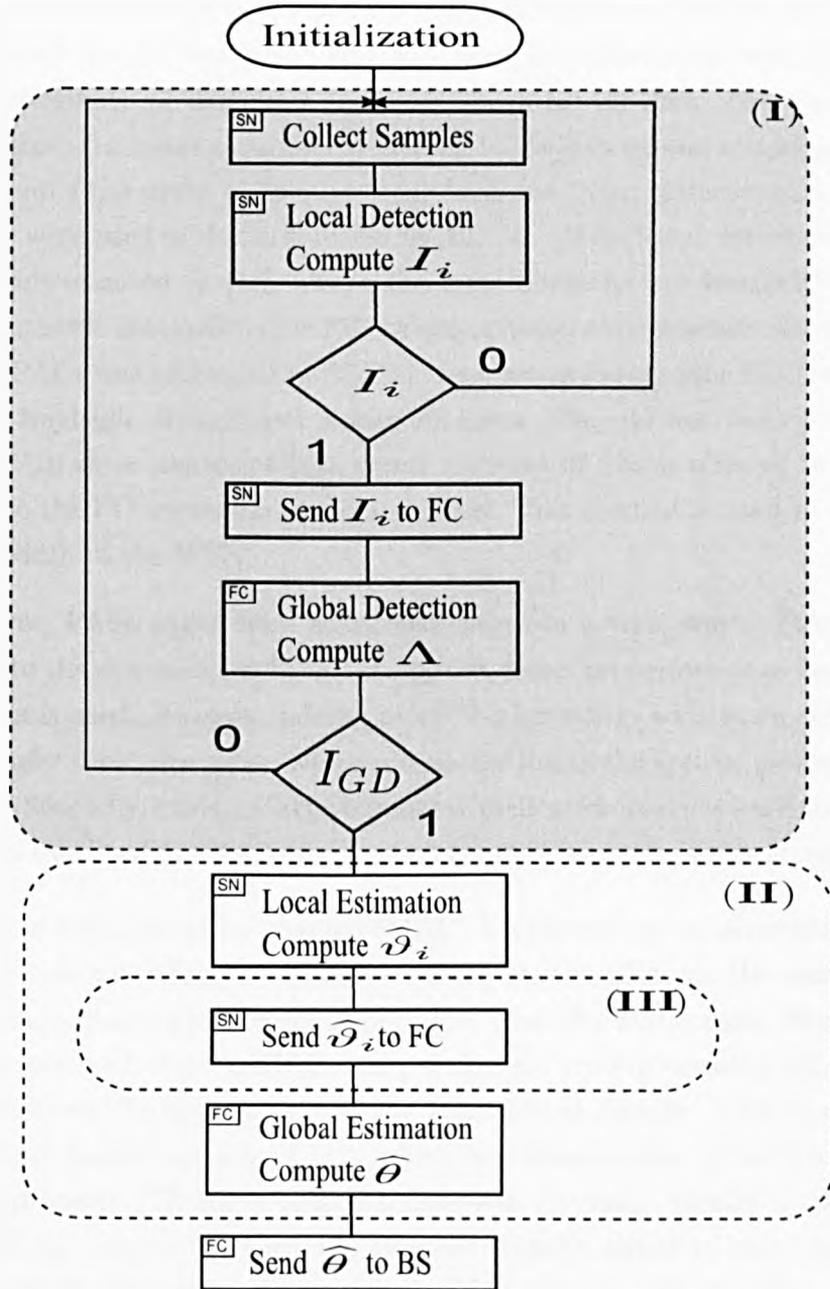


Figure 2.1.1: Distributed Detection, Localization and Estimation (DDLE) algorithm. In this chapter, section (I) will be discussed.

2.1. INTRODUCTION

2.1.2 Related Work

There is a large body of literature studying the problem of decision fusion over ideal parallel access channels (PACs) [33, 34, 35, 36]. An optimum decision fusion rule over such a channel was derived in [62] and [72]. However both papers require knowledge of local probability of detection and false alarm for all SNs. Niu and Varshney [37, 38] relaxed the latter requirement leading to the suboptimal counting rule (CR), which is simply the count of positive local decisions. Scan statistics and a Bayesian framework were used in decision fusion in [73, 74]. Distributed detection in a noisy PAC was investigated in [41], where the local decisions are transmitted through binary symmetric channels to the FC. Decision fusion over coherent Rayleigh fading and noisy PACs was addressed in [75, 39]. Incoherent fading over PACs was studied in [76] for Rayleigh, Rician, and Nakagami cases. The authors there also consider “censoring” in their treatment [32], where a subset of SNs is allowed to relay their decisions to the FC according to some criteria. This method is used to save power and bandwidth in the WSN.

However, PACs suffer from scalability issues in several ways. Firstly, adding more SNs to the system is expected to improve detection performance because more information is used. However, adding more SNs implicitly adds more parallel channels that have their own noise, which ultimately limits the system performance, see Fig. 2.1.2. Secondly, having a large number of dedicated channels leads to large communication overhead and increased system complexity such as synchronization.

Hence, a multiple access channel (MAC) is proposed as an alternative because it is more resource efficient. Moreover, since all the SNs use the same channel, the communication noise is considerably less than the PACs case. From a signal processing point of view, a MAC can provide an array-processing gain since the receiver observes the aggregation of the transmitted signals. This is also known as *in-channel fusion*¹(see Fig. 2.1.3), which is a general case of the source-channel matching property [77] for a coherent Gaussian channel. Ideally, if the received signals add up coherently, then the received signal’s signal to noise ratio (SNR) will dramatically increase as the number of SNs tend to infinity. This observation has been exploited in [78] for the type-based multiple access (TBMA) regime and extended to coherent fading channels in [79]. Note however, that an independent channel is needed for every type used.

¹In-channel fusion was used in distributed estimation in [60].

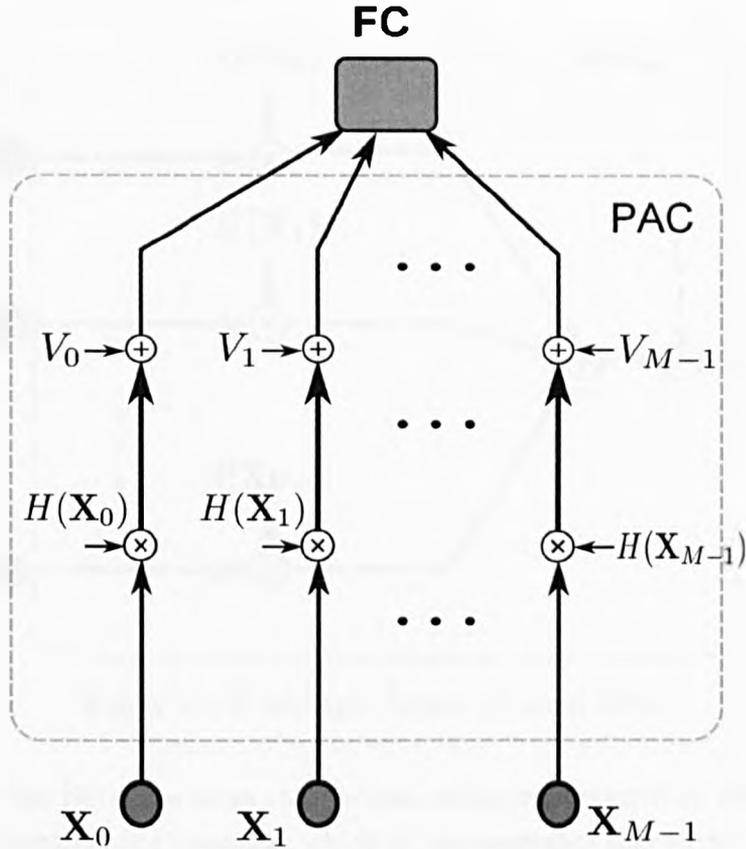


Figure 2.1.2: Parallel Access Channel WSN.

Incoherent Rayleigh and Rician fading MACs were considered in [80] where the authors suggested incoherent detection by using an ED. The authors showed that the error probability hits a floor even when increasing the number of SNs in the Rayleigh channel. To overcome this issue, the use of multiple antennas was suggested in [81]. Similarly, Diversity-MAC² was proposed in [82], where several MACs are used to provide diversity. In each MAC though, incoherent detection is performed via an ED.

2.1.3 Chapter Contributions

In this chapter, we investigate distributed detection of a target with unknown parameters in a censored WSN suffering from an incoherent Rayleigh fading³ MAC. Interestingly, the above treatment of distributed detection takes a similar course as in data communication systems, where conventional diversity combining techniques (such as maximum ratio combining and equal gain combining) are suggested. Although such an approach can tap into the vast literature of diversity combining, it

²We call it here centralized diversity combining (CDC) MAC.

³We refer to incoherent Rayleigh fading with random phase simply as Rayleigh fading.

2.1. INTRODUCTION

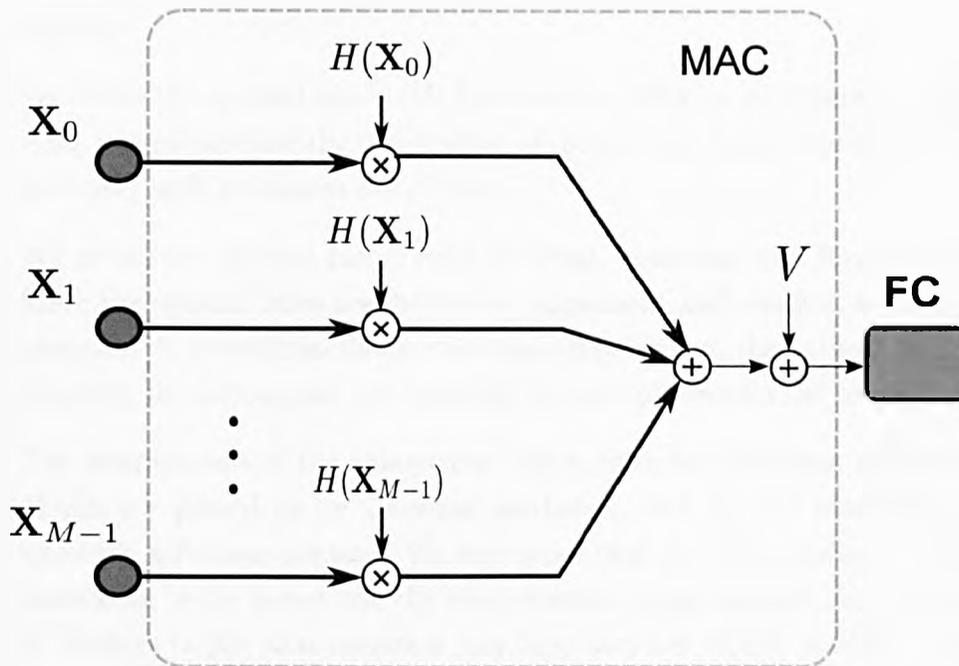


Figure 2.1.3: Multiple Access Channel WSN.

requires that the FC has a more complicated receiver compared to other SNs. This leads to a heterogeneous network, which is not desirable due to cost and robustness considerations. Hence, we take the next logical step of using *precoding* [83] to overcome the fading problem in distributed detection in WSNs. The performance of the discussed distributed detection algorithms is also analyzed using a stochastic geometry framework.

We gradually build up the discussion to reach the distributed diversity combining method. First, the optimal fusion rule over parallel access channels is discussed. By identifying its drawbacks, the counting rule is motivated here and is also considered over PACs. After that, distributed detection over MAC is shown to be more efficient than the PACs case. Hence, we discuss different channel assumptions starting with the ideal channel, then moving to the Gaussian MAC, and finally the Rayleigh fading MAC. In order to overcome the deterioration due to fading, the centralized diversity combining method [82] is revised. Finally, inspired by precoding, the simpler more efficient distributed diversity combining methods are suggested. We summarize our contributions in the following:

- We show that the detecting SNs under the alternative hypothesis form an inhomogeneous Cox PPP, whereas they form a homogeneous PPP under the null hypothesis. Consequently, the number of detecting SNs follows a Poisson mixture and Poisson distribution under the alternative and null hypothesis respectively. This is contrary to the results in [41] that use Poisson approxi-

2.1. INTRODUCTION

mation.

- We derive the optimal fusion rule for a random network with parallel channels. Also, we approximate the distribution of the optimal fusion rule using moment matching with a Gamma distribution.
- We derive the optimal fusion rules for ideal, Gaussian, and Rayleigh MACs. Since the optimal rules are difficult to implement and analyze, we derive implementable suboptimal fusion rules that upper bound the optimal rules. Fortunately, the suboptimal rule is simply a threshold test on the received signal.
- The distributions of the suboptimal fusion rules for Gaussian, and Rayleigh MACs are proved to be Gaussian mixtures. But for the ideal MAC it is generally a Poisson mixture. We also prove that the CR is optimal under full knowledge of the target and the ideal channel in any network size, in contrast to findings in [38] that require a very large number of SNs in order to invoke the central limit theorem.
- We propose two distributed diversity combining methods, inspired from the precoding concept, to mitigate the fading effects. The distributed maximum ratio combining (dMRC) and distributed equal gain combining (dEGC). The dMRC and dEGC are in fact suboptimal solutions to the problem of maximizing the J-divergence for distributed detection. We also show that distributed diversity combining is able to overcome the performance floor reported in [84]. Furthermore, it performs better than the limitation of using CDC MAC [82] under low and medium SNRs. Better performance is attained by increasing SNs deployment density and optimizing the local detection threshold.

2.1.4 Chapter Outline

This chapter is organized as follows. A brief introduction about stochastic geometry and the system model are provided in Section 2.2. In Section 2.3, we derive statistical properties of the SN point process. Those results are then used to characterize the performance of the distributed detection algorithms. Section 2.4 provides the discussion of distributed detection of PAC that paves the way for Section 2.5, which investigates fusion rules over MAC under different channel assumptions. In Section 2.6 we propose the distributed diversity combining techniques and derive the detection performance for the various diversity combining techniques. Simulation results are also included in this section. Finally we conclude the chapter with Section 4.6.

2.2. PROBLEM FORMULATION

2.2 Problem Formulation

The problem is formulated by first modeling the target and the WSN. Then the distributed detection set up is considered.

2.2.1 System Model

We model the elements of the distributed detection system in addition to the target. Related assumption are provided as well.

Target

We assume that an intruder target is a point source randomly located at \mathbf{X}_t . The target emits energy that propagates in a spherical wave. Consequently, the received energy at any SN is inversely proportional to the square of the distance separating the target and the SN. However, having a non-ideal medium leads to further loss in energy. Hence, the received energy decay can be confidently modeled as an inverse power law with exponent $e \geq 1$. This model also reasonably describes electromagnetic or acoustic sources. While, the target's amplitude $A > 0$ is usually unknown, we will assume that it follows a known probability distribution $p_A(a)$.

Sensor node

SNs are deployed in a sensing field \mathcal{F} where $\mathcal{F} \subset \mathbb{R}^2$. They are located at points $\{\mathbf{X}_i\}$ that are independent and identical distributed (i.i.d) according to a uniform distribution. SNs are primarily equipped with a physical sensing device, local detector, information extraction unit, and a transceiver. Each SN samples the physical medium and collects an N even number⁴ at a sampling frequency of $1/T_s$. For example, the SN might be equipped with a pyrometric sensor to measure infrared energy [85]. The measurements are corrupted by i.i.d additive white Gaussian noise in space and time having zero mean and known variance σ_s^2 .

Communication channel

The communication channel is shared and assumed to suffer from an AWGN noise having the distribution $\mathcal{N}(0, \sigma_c^2)$, where σ_c^2 is assumed to be known. Moreover, the

⁴Having an even number of samples makes the Marcum Q-function in (2.2.9) tractable.

2.2. PROBLEM FORMULATION

transmissions between the detecting SNs and the FC occurs over an i.i.d Rayleigh fading channels with gain $H(\mathbf{X}_i) \sim \mathcal{CN}(0, \sigma_h^2)$ between SN at \mathbf{X}_i and the FC, where σ_h^2 is the channel power. Note that the channel $H(\mathbf{X}_i)$ does not depend on the SN's location \mathbf{X}_i in contrast to what might the notation here imply.

WSN

The WSN consists of a fusion center connected to a group of SNs dispersed over the sensing field. The FC task is twofold: fusing the received local decisions into global detection decision and broadcasting a periodic pilot signal, which is used for channel state information (CSI) estimation and synchronization by the SNs.

We use the elegant framework of stochastic geometry to model the previous WSN as a stationary marked Poisson point process (MPPP) $\Phi = \{(\mathbf{X}_i, \Xi(\mathbf{X}_i), H(\mathbf{X}_i))\}$. The MPPP represents the SN as *points* located at \mathbf{X}_i uniformly distributed in \mathcal{F} . The SNs carry two marks, the first is the local test statistic $\Xi(\mathbf{X}_i)$ and the second is the channel gain $H(\mathbf{X}_i)$ between the SN and the FC. Finally, we make the plausible assumption of the target not being co-locating with any of the SNs in the network, i.e., $\mathbf{X}_t \notin \Phi$.

2.2.2 Distributed Detection

2.2.2.1 Local Detection

Upon taking measurements, a SN located at \mathbf{X}_i tests for one of the following hypotheses

$$\mathcal{H}_0 : F_i[n] = W_i[n] \quad (2.2.1)$$

$$\mathcal{H}_1 : F_i[n] = \frac{A}{\|\mathbf{X}_i - \mathbf{X}_t\|^e} + W_i[n] \quad (2.2.2)$$

where $n = 0, \dots, N - 1$. The SN processes the measurements through a local detector to produce a local test statistic $\Xi(\mathbf{X}_i)$, which is distributed according to $p_{\Xi}(\xi)$. The local detector here can take any form, such as an ED for example. The local detection decision is found via a threshold test on the local test statistic, i.e.,

$$I(\mathbf{X}_i) = \begin{cases} 1, & \Xi(\mathbf{X}_i) \geq \gamma \\ 0, & \Xi(\mathbf{X}_i) < \gamma \end{cases} \quad (2.2.3)$$

2.2. PROBLEM FORMULATION

where γ is the local detection threshold. Now, $I(\mathbf{X}_i)$ is a binary RV characterized by the false alarm probability

$$P_{fa} = \mathbb{P}(\Xi(\mathbf{X}_i) \geq \gamma; \mathcal{H}_0) \quad (2.2.4)$$

and the detection probability

$$P_d = \mathbb{P}(\Xi(\mathbf{X}_i) \geq \gamma; \mathcal{H}_1). \quad (2.2.5)$$

Target with known parameters

If the target's parameters are known at the SNs, the optimal detector would be the matched filter (MF). In this case, the false alarm probability of the MF is given by

$$P_{fa}^{\text{MF}} = Q(\gamma) \quad (2.2.6)$$

where $Q(\cdot)$ is the Gaussian Q-function with $Q(z) = \int_z^\infty \exp\left(-\frac{y^2}{2}\right) dy$. Whereas for a given target amplitude (a) located at \mathbf{x}_t and SN located at \mathbf{x}_i , the detection probabilities for the MF is then given by

$$P_d^{\text{MF}}(\mathbf{x}_i, a) = Q\left(\gamma - \frac{a}{\sigma_s \|\mathbf{x}_i - \mathbf{x}_t\|}\right). \quad (2.2.7)$$

Unknown target parameters

On the other hand, if the target's parameters are unknown, which is the case in passive detection applications, the MF cannot be used. A suitable alternative would be the ED since it does not need any prior knowledge about the target. The false alarm probability of the ED is given by [86]

$$P_{fa}^{\text{ED}} = Q_{\chi_N^2}(\gamma) \quad (2.2.8)$$

where $Q_{\chi_N^2}(\cdot)$ is the tail probability of the central χ^2 distribution with N degrees of freedom. The detection probability, on the other hand, is

2.2. PROBLEM FORMULATION

$$P_d^{\text{ED}}(\mathbf{x}_i, a) = Q_{\frac{N}{2}}\left(\frac{a}{\sigma_s \|\mathbf{x}_i - \mathbf{x}_t\|}, \sqrt{\gamma}\right) \quad (2.2.9)$$

where $Q_{\frac{N}{2}}(\cdot, \cdot)$ is the generalized Marcum Q-function [86]. N here should be an even number in order for the Marcum Q-function to be tractable.

However, as seen in (2.2.9), the target's parameters are still needed to find the detection probability. Also those parameters are needed to find the global detection performance as we shall show later. One way to handle this matter is via estimating the required parameters, but this would lead to complications in the performance analysis. Another way is to assign a probability distribution to the parameters and use the Bayesian approach. The resulting theoretical performance would bound the actual performance. In this chapter we deal with the unknown parameter problem in two stages. The first eliminates the dependance on the target location using the stochastic geometry framework and the second uses the Bayesian framework to eliminate the dependance on its amplitude.

2.2.2.2 Global Detection

The SNs collaborate together to reach the global decision about the target. They send their local decisions to the FC for global processing. The overall system is shown in Fig. 2.2.1.

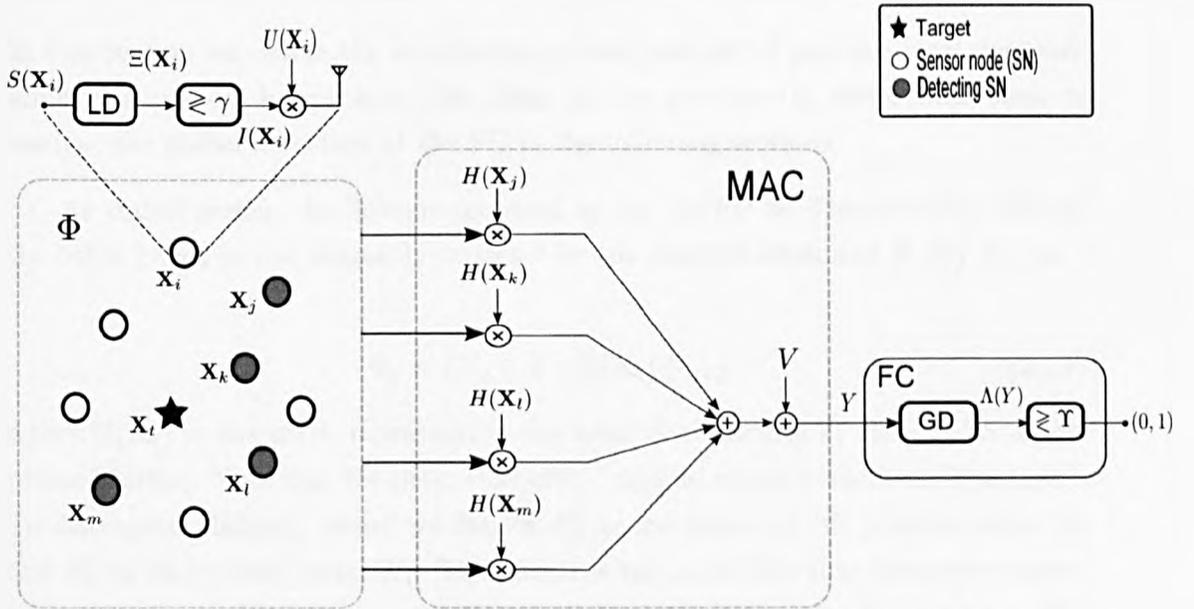


Figure 2.2.1: Distributed Detection in Wireless Sensor Networks System Diagram.

2.3. STOCHASTIC GEOMETRY ANALYSIS OF DETECTING SENSOR NODES

To conserve energy, the SNs are “censored”, i.e., $I(\mathbf{X}_i)$ is sent to the FC only if it is a positive decision. In other words, the SNs use a form of on-off keying (OOK) modulation for transmission, i.e., $U(\mathbf{X}_i) = cI(\mathbf{X}_i)$, where c is an arbitrary value chosen to improve the global detection performance, with further discussion given later in Subsection 2.6. Let Y be the received signal at the FC, which is given by

$$Y = \sum_{\mathbf{X}_i \in \Phi} H(\mathbf{X}_i)U(\mathbf{X}_i) + V \quad (2.2.10)$$

where V is the AWGN. The FC process the received signal via a global detector, $\Lambda(\cdot)$, to reach the global decision

$$I_{GD} = \begin{cases} 1, & \Lambda(Y) \geq \Upsilon \\ 0, & \Lambda(Y) < \Upsilon. \end{cases} \quad (2.2.11)$$

Given the model uncertainties, σ_s^2 from the sensing side, in addition to σ_c^2 and σ_h^2 from the communication side, we wish to design efficient local and global processing algorithms, $U(\mathbf{X}_i)$ and $\Lambda(Y)$ respectively⁵.

2.3 Stochastic Geometry Analysis of Detecting Sensor Nodes

In this section we derive the distribution of the number of positive local decisions under \mathcal{H}_0 and \mathcal{H}_1 hypotheses. By doing so, we provide the theoretical basis to analyze the global detection at the FC in the following sections.

As stated earlier, the SNs are modeled by the MPPP Φ . The detecting SNs on the other hand, can be elegantly modeled by the *thinned* version of Φ , say Φ_d , as

$$\Phi_d = \{\mathbf{X}_i \in \Phi : \Xi(\mathbf{X}_i) \geq \gamma\} \quad (2.3.1)$$

where $\Xi(\mathbf{X}_i)$ is the mark representing the local test statistic at the i th SN as explained earlier. Note that the properties of Φ_d , depend on the considered hypothesis. To distinguish different cases, we denote Φ_d^0 as the detecting SN process under \mathcal{H}_0 and Φ_d^1 as the process under \mathcal{H}_1 . The former is the set of SNs that have experienced false detection and the latter set is the set of detecting SNs under \mathcal{H}_1 . Let $\Omega_j = |\Phi_d^j|$ for $j = 0, 1$ be the number of SNs in Φ_d^j or the cardinality of Φ_d^j , which is a quantity

⁵ $U(\mathbf{X}_i)$ can be viewed as a modulation scheme as in [82].

2.3. STOCHASTIC GEOMETRY ANALYSIS OF DETECTING SENSOR NODES

that provides significant information about the detecting PP. It follows that under \mathcal{H}_j

$$\Omega_j = \sum_{\mathbf{x}_i \in \Phi} I(\mathbf{X}_i) | \mathcal{H}_j, j = 0, 1. \quad (2.3.2)$$

Obviously Ω_j is a RV whose distribution is dependent on Φ_d^j properties. To statistically characterize Φ_d^j , we present the following theorem.

Theorem 2.1. *The detecting SNs under hypothesis \mathcal{H}_0 form a homogeneous MPPP Φ_d^0 , and so Ω_0 follows the Poisson distribution⁶*

$$p_0(\omega) = \text{Poi}(\bar{\Omega}_0) \quad (2.3.3)$$

where

$$\bar{\Omega}_0 = \lambda P_{fa} |\mathcal{F}| \quad (2.3.4)$$

is the mean of Ω_0 and $|\mathcal{F}|$ is the area of the sensing field.

Proof. See Appendix A.1. □

The above theorem says that the active SNs due to constant false alarm are uniformly distributed in the sensing field with intensity $\bar{\Omega}_0$, since the point process of active SNs is homogenous. It also says that the distribution is independent of the local detector used because γ can be set arbitrarily.

Under \mathcal{H}_1 however, the distribution of Ω_1 is rather more challenging, due to having different local detection probability at each SN. Nevertheless, we can describe Ω_1 's distribution using the stochastic geometry framework set up in Section 2.2. Thus, we present our main result in the following theorem.

Theorem 2.2. *It is assumed that there exist a stationary homogenous MPPP Φ and an arbitrary amplitude distribution for the target, say $p_A(a)$. Furthermore, assume that \mathcal{F} is infinitely large. Then, under hypothesis \mathcal{H}_1 for any local detector the detecting SNs constitute an inhomogeneous Cox⁷ MPPP Φ_d^1 . Consequently, the number of detecting SNs, Ω_1 , has the distribution*

$$p_1(\omega) = \mathbb{E}_A [\text{Poi}(\Delta(A))] \quad (2.3.5)$$

⁶Equation (2.3.3) actually means $p_0(\omega) = (\bar{\Omega}_0^\omega / \omega!) \exp(-\bar{\Omega}_0)$.

⁷Which is also known as mixed.

2.3. STOCHASTIC GEOMETRY ANALYSIS OF DETECTING SENSOR NODES

where

$$\Delta(A) = \lambda \int_{\mathcal{F}} P_d(\mathbf{x}, A) d\mathbf{x}. \quad (2.3.6)$$

It follows also that the mean is

$$\bar{\Omega}_1 = \lambda \mathbb{E}_A \left[\int_{\mathcal{F}} P_d(\mathbf{x}, A) d\mathbf{x} \right]. \quad (2.3.7)$$

Proof. See Appendix A.2. □

The theorem states that the distribution of the detecting SNs using *any* local detector of *any* target's amplitude distribution follows an inhomogeneous Cox Poisson distribution with mean intensity depending on the local detector's probability of detection. The spatial distribution of the detecting SNs, in contrast to theorem 2.1, is concentrated around the target and gradually decays as we move further away. It is noted from the proof that the distribution does not depend on the target's location if \mathcal{F} is sufficiently large.

A special case of theorem 2.2 is when the target's amplitude is deterministic. Then, the distribution simplifies to Poisson as stated in the following corollary.

Corollary 2.1. *If the target's amplitude is deterministic (say a), then the detecting SNs constitute an inhomogeneous PPP, and so Ω_1 follows the following Poisson distribution*

$$p_1(\omega) = \text{Poi}(\bar{\Omega}_1) \quad (2.3.8)$$

where

$$\bar{\Omega}_1 = \lambda \int_{\mathcal{F}} P_d(\mathbf{x}, a) d\mathbf{x}. \quad (2.3.9)$$

Furthermore, the above corollary applies to the case when the target signal is space-time varying one such as

$$s(\mathbf{x}, t) = \frac{c}{\|\mathbf{x} - \mathbf{x}_t\|} \text{erfc} \left(\frac{\|\mathbf{x} - \mathbf{x}_t\|}{2\sqrt{\kappa t}} \right), \quad t > 0 \quad (2.3.10)$$

where c and κ are some constants and $\text{erfc}(\cdot)$ is the complementary error function. The above represents a diffusion signal generated by a source at \mathbf{x}_t , which is the

2.3. STOCHASTIC GEOMETRY ANALYSIS OF DETECTING SENSOR NODES

model used in Chapter 3. In this case, the detecting SNs distribution adheres to (2.3.8) with

$$\bar{\Omega}_1 = \lambda \int_{\mathcal{F}} Q_{\frac{N}{2}} \left(\sqrt{\frac{a_s^2(\mathbf{x})}{\sigma_s^2}}, \sqrt{\gamma} \right) d\mathbf{x} \quad (2.3.11)$$

if the ED is used where

$$a_s(\mathbf{x}) = \sqrt{\frac{c^2}{\|\mathbf{x} - \mathbf{x}_t\|^2} \sum_{n=0}^{N-1} \operatorname{erfc}^2 \left(\frac{\|\mathbf{x} - \mathbf{x}_t\|}{2\sqrt{\kappa n T_s}} \right)}. \quad (2.3.12)$$

A similar analysis applies for the MF detector. Note however, that the local detector averages (in some sense) the time signal leaving only the spatial aspect of the signal. Thus, our analysis in this chapter applies for space-time signal models as well without any significant change in analysis. However the power law model is simpler in its treatment.

Returning back to the power law model, if the target's amplitude is random, finding Ω_1 's distribution is not trivial since it requires the expectation with respect to (w.r.t) the distribution of Δ in (2.3.6), which in turn depends on the distribution of A through the local detector used. Indeed, Δ is a transformation of the random amplitude of the target A , i.e., equation (2.3.6) can be simply interpreted as $\Delta = g(A)$. Consequently, Δ 's distribution is a function of $p_A(a)$. Nonetheless, we cannot find the distribution using $g(\cdot)$ directly since it is not analytically tractable. To address this issue we approximate $g(\cdot)$ with a polynomial function that can be easily inverted to find Δ 's distribution. To do that, we show that $\mathbb{E}_A [P_d(\mathbf{x}, A)]$ can be approximated by a piecewise constant function. In fact, this approximation is known in the literature as the *disk model* [87]. We adopt a similar approximation in which the SNs detect a target with probability one within a certain range and experience false detection (if any) fall beyond that range. This claim is reinforced by the fact that the signal's power decays as the distance from the target increases, see Fig. 2.3.1. Hence, there exists a particular range after which the target's signal diminishes to approximately zero. Hence, if any SN beyond this range reports a detection, it is due a false alarm. Therefore, $\Delta(a)$ can be approximated by a piecewise quadratic function, since P_d depends on the signal's SNR.

In order to prove the previous statement formally, first we use the J-divergence as a measure to distinguish between \mathcal{H}_1 and \mathcal{H}_0 given a set of measurements at an arbitrary SN. Recall that the J-divergence [88] between distributions $p_0(z)$ and $p_1(z)$, which is defined as

2.3. STOCHASTIC GEOMETRY ANALYSIS OF DETECTING SENSOR NODES

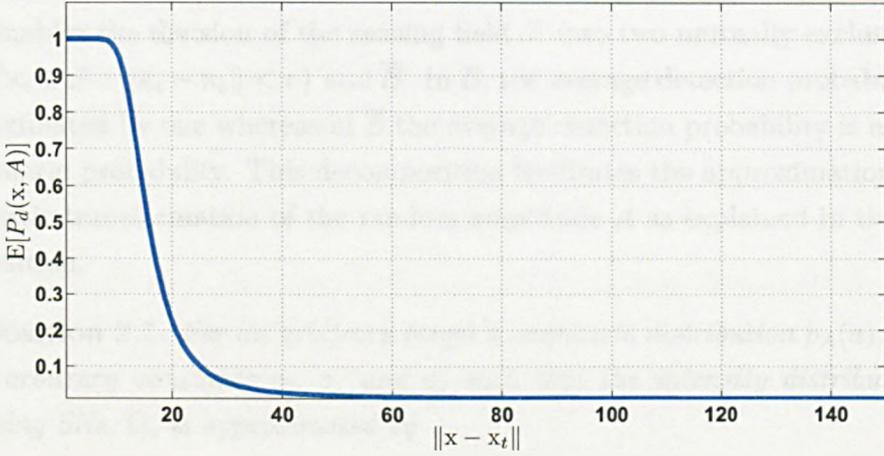


Figure 2.3.1: The averaged local probability of detection ($\mathbb{E}_A [P_d(\mathbf{x}, A)]$) plotted against the distance from the target ($\|\mathbf{x} - \mathbf{x}_t\|$).

$$J(p_0(z), p_1(z)) = \mathbb{E}_1 \left[\log \left(\frac{p_1(z)}{p_0(z)} \right) \right] - \mathbb{E}_0 \left[\log \left(\frac{p_1(z)}{p_0(z)} \right) \right] \quad (2.3.13)$$

where $\mathbb{E}_1[\cdot]$ and $\mathbb{E}_0[\cdot]$ are the expectations w.r.t the distributions $p_1(z)$ and $p_0(z)$. Next we bound the J-divergence between the local test statistic distribution under the two hypotheses in the following lemma.

Lemma 2.1. *For an arbitrary SN at \mathbf{x}_i and target located at \mathbf{x}_t , let an arbitrary test statistic be $\Xi_i = \Xi(\mathbf{X}_i)$, with realization ξ_i . Also let $e = 1$ (in (2.2.2)) without loss of generality, then J-divergence of the measured data's distribution under \mathcal{H}_0 and \mathcal{H}_1 satisfies*

$$J(p_0(\xi_i), p_1(\xi_i)) \leq \frac{N\mathbb{E}[A^2]}{\sigma_s^2 \|\mathbf{x}_i - \mathbf{x}_t\|^2}.$$

Proof. See Appendix A.3. □

Thus, for an arbitrary $\epsilon > 0$, there exists an SN located at some distance $r > 0$ away from the target such that

$$J(p_0(\xi_i), p_1(\xi_i)) \leq \frac{N\mathbb{E}[A^2]}{\sigma_s^2 r^2} < \epsilon. \quad (2.3.14)$$

In other words, given a target at \mathbf{x}_t , we can find SNs located at $\mathbf{x}_i \in \mathcal{F}$ for which the J-divergence is arbitrarily small, i.e., SNs cannot discriminate between \mathcal{H}_1 and \mathcal{H}_0 . This means that later SNs do not actually sense the target. Con-

2.4. DECISION FUSION OVER IDEAL PARALLEL CHANNELS

sequently, the *useful* SNs are located at $\{\mathbf{x}_i \in \Phi : \|\mathbf{x}_i - \mathbf{x}_t\| < r\}$. This observation enables the division of the sensing field \mathcal{F} into two mutually exclusive regions $\mathcal{B} = \{\mathbf{x}_i \in \mathcal{F} : \|\mathbf{x}_i - \mathbf{x}_t\| < r\}$ and $\bar{\mathcal{B}}$. In \mathcal{B} , the average detection probability can be approximated by one whereas in $\bar{\mathcal{B}}$ the average detection probability is actually the false alarm probability. This decomposition facilitates the approximation of Δ as a quadratic transformation of the random amplitude A as explained in the following proposition.

Proposition 2.1. *For an arbitrary target's amplitude distribution $p_A(a)$, then there exist arbitrary constants c_0 , c_1 and c_2 such that the intensity distribution of the detecting SNs, Ω , is approximated by*

$$p_\Omega(\omega) \approx \frac{1}{2\sqrt{c_0(\omega - c_1)}} f_A \left(c_2 + \sqrt{\frac{\omega - c_1}{c_0}} \right) + \frac{1}{2\sqrt{c_0(\omega - c_1)}} f_A \left(c_2 - \sqrt{\frac{\omega - c_1}{c_0}} \right) \quad (2.3.15)$$

Proof. See Appendix A.4. □

The proposition states that $\Delta \approx c_0(A - c_1)^2 + c_2$, and as a result, an inverse relation can be readily attained and consequently so can the distribution of Δ given an arbitrary distribution for A . However, the quadratic coefficients are dependent on the local detector. Incorporating the above proposition with Theorem 2.2, the distribution of the number of detecting SNs is completely characterized under \mathcal{H}_1 .

Remark 2.1. We digress here for a moment and revisit Theorem 2.2 in the light of the above observations. Recall that \mathcal{F} is assumed to be infinitely large or equivalently the target is assumed to be in the center of \mathcal{F} . Although the previous condition simplifies the proof of the theorem, it is rather strict. Using the previous argument, this condition can be relaxed to having the target at r units away from the edge of \mathcal{F} .

2.4 Decision Fusion over Ideal Parallel Channels

In this section we discuss some of the fusion rules used in parallel channels. The aim of this section is to first explore the difficulties in the parallel channel decision fusion in order to set up the stage for the MAC case. The optimal fusion rule (OFR) for parallel channels is discussed first. Then the suboptimal counting rule is proposed as an alternative.

2.4. DECISION FUSION OVER IDEAL PARALLEL CHANNELS

2.4.1 Optimal Decision Fusion Rule

Although this scenario is far from realistic in terms of assumptions and resources required, it gives an upper bound for distributed detection algorithms. The way this matter is tackled is by means of Neyman-Pearson's approach, i.e., by formulating the likelihood ratio test. We proceed as follows.

Given a PPP Φ , the optimal fusion of independent local decision is attained through the likelihood ratio [57]

$$\Lambda = \prod_{\mathbf{x}_i \in \Phi} \frac{p_{I(\mathbf{x}_i)}(d; \mathcal{H}_1)}{p_{I(\mathbf{x}_i)}(d; \mathcal{H}_0)} \quad (2.4.1)$$

where $p_{I(\mathbf{x}_i)}(d; \mathcal{H}_j)$ for $j = 0, 1$ is the distribution of the local decision of the SN located at \mathbf{X}_i , $I(\mathbf{X}_i)$, given the corresponding hypothesis. Note that in contrast to algorithms in the sequel sections, all SNs are required to report their local decision to the FC. Under both hypotheses however, $I(\mathbf{X}_i)$ obviously takes one of two values: a one for positive detection and a zero for negative decision, which leads to the local decision's distribution:

$$p_{I(\mathbf{x}_i)}(d; \mathcal{H}_1) = P_d(\mathbf{X}_i)I(\mathbf{X}_i) - (1 - P_d(\mathbf{X}_i))(1 - I(\mathbf{X}_i)) \quad (2.4.2)$$

$$p_{I(\mathbf{x}_i)}(d; \mathcal{H}_0) = (1 - P_{fa})I(\mathbf{X}_i) - P_{fa}(1 - I(\mathbf{X}_i)). \quad (2.4.3)$$

For simplicity, we assume that the local detector is a MF, and hence the target's parameters are assumed to be known. The above equations imply an interesting fact about distributed detection. Obviously the SNs can be from the detecting SN set or its complement, i.e., $\mathbf{X}_i \in \Phi_d$ or $\mathbf{X}_i \in \bar{\Phi}_d$. However, it is an element in Φ_d if it detects the target if it exists or falsely detects it if the target is absent. Similarly, it is in $\bar{\Phi}_d$ if it misdetects the target if it exist or it does not detect it if it is actually absent. Hence, (2.4.1) can be decomposed into

$$\begin{aligned} \Lambda &= \prod_{\mathbf{x}_i \in \Phi_d} \frac{p_{I(\mathbf{x}_i)}(d; \mathcal{H}_1)}{p_{I(\mathbf{x}_i)}(d; \mathcal{H}_0)} \prod_{\mathbf{x}_i \in \bar{\Phi}_d} \frac{p_{I(\mathbf{x}_i)}(d; \mathcal{H}_1)}{p_{I(\mathbf{x}_i)}(d; \mathcal{H}_0)} \\ &= \prod_{\mathbf{x}_i \in \Phi_d} \frac{P_d(\mathbf{X}_i)}{P_{fa}} \prod_{\mathbf{x}_i \in \bar{\Phi}_d} \frac{1 - P_d(\mathbf{X}_i)}{1 - P_{fa}} \end{aligned} \quad (2.4.4)$$

where the first product represents the likelihood ratio of the detecting SNs, whereas

2.4. DECISION FUSION OVER IDEAL PARALLEL CHANNELS

the second product represents the non-detecting SN's likelihood ratio. Taking the logarithm gives the log-likelihood ratio (LLR):

$$\Lambda_{\text{OFR}} = \log \Lambda = \underbrace{\sum_{\mathbf{X}_i \in \Phi_d} \log \frac{P_d(\mathbf{X}_i)}{P_{fa}}}_{\Lambda_{\text{OFR}}^1} + \underbrace{\sum_{\mathbf{X}_i \in \bar{\Phi}_d} \log \frac{1 - P_d(\mathbf{X}_i)}{1 - P_{fa}}}_{\Lambda_{\text{OFR}}^0} \quad (2.4.5)$$

which is similar to the Chair-Varshney rule [62]. However, the distribution of the LLR is rather complicated. Its exact distribution is the M -fold convolution of the local log-likelihood ratios distributions of all the SNs [72], where M is a RV. Hence, the LLR in (2.4.5) is difficult to obtain in practice. Therefore, we resort to approximating it by moment matching with either a Gaussian distribution or a Gamma distribution [66], which has the following form:

$$\text{Gamma}(\alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\beta} \quad (2.4.6)$$

where $\Gamma(\cdot)$ is the Gamma function, and α and β are the shape and scale parameters respectively. Moment matching method is used here since we are able to find the first two moments analytically, which generally specify the statistical behavior. Thus, it is proposed to match those moments with that of a known distribution such as the Gaussian or Gamma. We start by finding the mean and variance of the Λ_{OFR} , which are

$$\mathbb{E}[\Lambda_{\text{OFR}}] = \mathbb{E}[\Lambda_{\text{OFR}}^1] + \mathbb{E}[\Lambda_{\text{OFR}}^0] \quad (2.4.7)$$

$$\text{var}[\Lambda_{\text{OFR}}] = \text{var}[\Lambda_{\text{OFR}}^1] + \text{var}[\Lambda_{\text{OFR}}^0]. \quad (2.4.8)$$

So we require the calculation of the four terms on the right hand side of the above equations. Before proceeding, we introduce the following functions and constants, which prove useful in finding the moments.

2.4. DECISION FUSION OVER IDEAL PARALLEL CHANNELS

$$f_0(\mathbf{x}) = \log \left(\frac{1 - P_d(\mathbf{x})}{1 - P_{fa}} \right) \quad (2.4.9)$$

$$f_1(\mathbf{x}) = \log \left(\frac{P_d(\mathbf{x})}{P_{fa}} \right) \quad (2.4.10)$$

$$\mathbb{I}_1 = \lambda \int_{\mathcal{F}} \log(f_0(\mathbf{x})) d\mathbf{x} \quad (2.4.11)$$

$$\mathbb{I}_2 = \lambda \int_{\mathcal{F}} \log^2(f_0(\mathbf{x})) d\mathbf{x} \quad (2.4.12)$$

$$\mathbb{I}_3 = \lambda \int_{\mathcal{F}} \log(f_1(\mathbf{x})) d\mathbf{x} \quad (2.4.13)$$

$$\mathbb{I}_4 = \lambda \int_{\mathcal{F}} \log^2(f_1(\mathbf{x})) d\mathbf{x} \quad (2.4.14)$$

where the \log is for base e . The required moments are computed in the following proposition.

Proposition 2.2. *If $P_d(\mathbf{x}) \geq P_{fa}$ for all \mathbf{x} , the Λ_{OFR}^1 and Λ_{OFR}^0 , under \mathcal{H}_1 , have the following means and variances:*

$$\mathbb{E} [\Lambda_{OFR}^1] = \bar{\Omega}_1 \mathbb{E}_1 [\log(f_1(x))] \quad (2.4.15)$$

$$\mathbb{E} [\Lambda_{OFR}^0] = \mathbb{I}_1 - \bar{\Omega}_1 \mathbb{E}_1 [\log(f_0(x))] \quad (2.4.16)$$

$$\text{var} [\Lambda_{OFR}^1] = \bar{\Omega}_1 \mathbb{E}_1 [\log^2(f_1(x))] \quad (2.4.17)$$

$$\text{var} [\Lambda_{OFR}^0] = \mathbb{I}_2 - \bar{\Omega}_1 \mathbb{E}_1 [\log^2(f_0(x))]. \quad (2.4.18)$$

Under \mathcal{H}_0 though, the mean and variance are

$$\mathbb{E} [\Lambda_{OFR}^1] = P_{fa} \mathbb{I}_3 \quad (2.4.19)$$

$$\mathbb{E} [\Lambda_{OFR}^0] = (1 - P_{fa}) \mathbb{I}_1 \quad (2.4.20)$$

$$\text{var} [\Lambda_{OFR}^1] = P_{fa} \mathbb{I}_4 \quad (2.4.21)$$

$$\text{var} [\Lambda_{OFR}^0] = (1 - P_{fa}) \mathbb{I}_3 \quad (2.4.22)$$

Proof. See Appendix A.5 □

The moments under \mathcal{H}_1 show a direct dependence on the average number of detecting SNs, whereas under \mathcal{H}_0 , the dependence is only on P_{fa} . Ultimately, the OFR mean is

2.4. DECISION FUSION OVER IDEAL PARALLEL CHANNELS

$$\mathbb{E}_1[\Lambda_{\text{OFR}}] = \mathbb{I}_1 + \bar{\Omega}_1 (\mathbb{E}_1[\log(f_1(x))] - \mathbb{E}_1[\log(f_0(x))]) \quad (2.4.23)$$

$$\mathbb{E}_0[\Lambda_{\text{OFR}}] = P_{fa}\mathbb{I}_3 + (1 - P_{fa})\mathbb{I}_1 \quad (2.4.24)$$

and the variance is

$$\text{var}_1(\Lambda_{\text{OFR}}) = \mathbb{I}_2 + \bar{\Omega}_1 (\mathbb{E}_1[\log^2(f_1(x))] - \mathbb{E}_1[\log^2(f_0(x))]) \quad (2.4.25)$$

$$\text{var}_0(\Lambda_{\text{OFR}}) = P_{fa}\mathbb{I}_4 + (1 - P_{fa})\mathbb{I}_3. \quad (2.4.26)$$

Now having the moments available, we are ready to perform moment matching. In the case of Gaussian distribution, the moments are plugged in directly. For the Gamma distribution case, the mean and variance of (2.4.6) are $\alpha\beta$ and $\alpha\beta^2$ respectively. It is easy to find that

$$\alpha_{\text{OFR},j} = \frac{\mathbb{E}_j[\Lambda_{\text{OFR}}]}{\text{var}_j(\Lambda_{\text{OFR}})} \quad (2.4.27)$$

$$\beta_{\text{OFR},j} = \frac{\text{var}_j(\Lambda_{\text{OFR}})}{(\alpha_{\text{OFR},j})^2} \quad (2.4.28)$$

for $j = 0, 1$ representing \mathcal{H}_0 and \mathcal{H}_1 cases.

Given the approximate distribution $\tilde{\Lambda}_{\text{OFR}} \sim \text{Gamma}(\alpha_{\text{OFR},j}, \beta_{\text{OFR},j})$ found by moment matching, the global detection probability P_D and the global false alarm probability P_{FA} can be found respectively as

$$P_{FA} \approx \mathbb{P}(\tilde{\Lambda}_{\text{OFR}} \geq \Upsilon; \mathcal{H}_0) \quad (2.4.29)$$

$$P_D \approx \mathbb{P}(\tilde{\Lambda}_{\text{OFR}} \geq \Upsilon; \mathcal{H}_1). \quad (2.4.30)$$

2.4.2 The Counting Rule

Unfortunately, the OFR requires the knowledge of all the detection probabilities in the WSN, which is very demanding. Hence, a suboptimal fusion rule was derived from (2.4.5) by Niu and Varshney known as the counting rule [37, 38]. In a random network, the CR simply takes the form of

2.5. DECISION FUSION OVER SHARED MAC

$$\Lambda_{\text{CR}} = \sum_{\mathbf{X}_i \in \Phi} I(\mathbf{X}_i) \quad (2.4.31)$$

where the local decisions are defined as $I = \{0, 1\}$.

The CR enjoys a few desirable features, such as implementation simplicity and no prior information about the local detection statistics. Furthermore, it promptly enables censoring. In other words, if the SN has a positive decision it sends to the FC otherwise it does not send anything. Yet another desirable feature is the ease of performance analysis. The performance is readily found by Theorems 2.1 and 2.2 (in subsection 2.3) for \mathcal{H}_0 and \mathcal{H}_1 cases respectively using the stochastic geometry framework. Thus, P_{FA} and P_D are given by

$$P_{FA} = \sum_{m=\Upsilon}^{\infty} \frac{\bar{\Omega}_0^m}{m!} e^{-\bar{\Omega}_0} \quad (2.4.32)$$

$$P_D = \sum_{m=\Upsilon}^{\infty} \mathbb{E}_A \left[\frac{\Delta^m(A)}{m!} e^{-\Delta(A)} \right] \quad (2.4.33)$$

where $\bar{\Omega}_0$ and $\Delta(A)$ are defined according to (2.3.4) and (2.3.6) respectively.

The theorems show that the distributions are exactly Poisson in contrast to [41] that uses Poisson for approximation. Using stochastic geometry also enables us to circumvent the issue of unknown target location, which was handled in [41] and [74] by integrating the target's position over the sensing field. This led to distributions described by multiple integrations that are even difficult to compute numerically. In contrast, our result is exact and numerically tractable.

2.5 Decision Fusion over Shared MAC

In this section we discuss fusion rules for a more resource efficient shared channel scenario. We start by investigating an ideal shared channel. Then we look into the Gaussian, and Rayleigh channel case. We start by investigating the optimal rules, which are, as it is usually the case, too complicated for analysis or implementation. Then we resort to suboptimal but simpler rules for the previous channels. In this section however, we adopt the concept of censoring [32] for the purposes of saving energy.

2.5. DECISION FUSION OVER SHARED MAC

2.5.1 Ideal Channel

Under the assumption of an ideal channel and perfect synchronization between the SNs and the FC. The received signal at the FC is

$$Y = \sum_{\mathbf{X}_i \in \Phi} I(\mathbf{X}_i). \quad (2.5.1)$$

The signal above is exactly the number of detecting SNs in (2.3.2), and therefore the distribution of Y is characterized by Theorems 2.1 and 2.2.

However, the distribution depends on the local detector used at the SN. The choice of a local detector depends on the information available about the target. On the one hand, we study the case in which the target's parameters are completely known. Although such assumption is strict from a practical point of view, it provides an upper bound on detection performance. On the other hand, we consider the case when only the statistics of the target's parameters are known, which we are more likely to encounter in practice.

Known Target Parameters

Assuming that the realizations a and \mathbf{x}_t of A and \mathbf{X}_t are available at the SN, the MF is used for local detection. The LLR at the FC given some realization \mathbf{y} of (2.5.1) is

$$\begin{aligned} \Lambda_{\text{I-MF}}(\mathbf{y}) &= \log \left(\frac{p_1(\mathbf{y})}{p_0(\mathbf{y})} \right) = \log \left(\frac{\text{Poi}(\bar{\Omega}_1)}{\text{Poi}(\bar{\Omega}_0)} \right) \\ &= \log \left(\frac{\bar{\Omega}_1^y e^{-\bar{\Omega}_1}}{\bar{\Omega}_0^y e^{-\bar{\Omega}_0}} \right) \\ &= y \log \left(\frac{\bar{\Omega}_1}{\bar{\Omega}_0} \right) + \bar{\Omega}_1 - \bar{\Omega}_0. \end{aligned} \quad (2.5.2)$$

Thus, the optimal fusion rule is $\Lambda_{\text{I-MF}}(\mathbf{y}) \geq \Upsilon$ which can be simplified into

$$\Lambda_{\text{I-MF}}(\mathbf{y}) = y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \Upsilon_1 \quad (2.5.3)$$

where

$$\Upsilon_1 = \frac{\Upsilon - \bar{\Omega}_1 + \bar{\Omega}_0}{\log(\bar{\Omega}_1) - \log(\bar{\Omega}_0)}.$$

2.5. DECISION FUSION OVER SHARED MAC

Note however, that (2.5.3) is actually the received signal itself. Interestingly, the CR, which was proposed as a suboptimal detector for parallel channels, turns out to be the optimal detector for the shared channel. The performance of the CR when using the MF as a local detector is fully described in corollary (2.1). Hence, the global false alarm probability is given by

$$P_{FA} = \sum_{m=\Upsilon_1}^{\infty} \frac{\bar{\Omega}_0^m}{m!} e^{-\bar{\Omega}_0} \quad (2.5.4)$$

and the global detection probability is

$$P_D = \sum_{m=\Upsilon_1}^{\infty} \frac{\bar{\Omega}_1^m}{m!} e^{-\bar{\Omega}_1}. \quad (2.5.5)$$

Unknown Target Parameters

If the target's parameters are unknown or actually random, we propose using the ED⁸ for local detection and assigning a prior distribution for the target's amplitude. The LLR follows as

$$\Lambda_{I-ED}(y) = \log \left(\frac{\mathbb{E}_A [\text{Poi}(\Delta(A))]}{\text{Poi}(\bar{\Omega}_0)} \right) \quad (2.5.6)$$

where in this case

$$\Delta(A) = \lambda \int_{\mathcal{F}} P_d^{\text{ED}}(\mathbf{x}, A) d\mathbf{x}. \quad (2.5.7)$$

Clearly, the optimal fusion rule (2.5.6) does not lend itself to performance analysis. Hence, a suboptimal (and simpler) fusion rule is formulated instead. Using Jensen's inequality to establish an upper bound on Λ_{I-ED} yields

⁸Since in practice it is unlikely that the target's parameters are known a-priori, the ED will be used for the rest of this chapter.

2.5. DECISION FUSION OVER SHARED MAC

$$\begin{aligned}\Lambda_{\text{I-ED}}(y) &\leq \tilde{\Lambda}_{\text{I-ED}}(y) \\ &= \mathbb{E}_A \left[\log \left(\frac{\text{Poi}(\Delta(A))}{\text{Poi}(\bar{\Omega}_0)} \right) \right]\end{aligned}\quad (2.5.8)$$

$$= \mathbb{E}_A \left[\log \left(\frac{\Delta^y(A) e^{-\Delta(A)}}{\bar{\Omega}_0^y e^{-\bar{\Omega}_0}} \right) \right]\quad (2.5.9)$$

$$\begin{aligned}&= y \left(\mathbb{E}_A [\Delta(A)] - \log(\bar{\Omega}_0) \right) \\ &+ \log(\bar{\Omega}_0) - \mathbb{E}_A [\log(\Omega(A))].\end{aligned}\quad (2.5.10)$$

Adopting $\tilde{\Lambda}_{\text{I-ED}}$ as our fusion rule, we again arrive at

$$\tilde{\Lambda}_{\text{I-ED}}(y) = y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\geq}} \Upsilon_2 \quad (2.5.11)$$

where

$$\Upsilon_2 = \frac{\Upsilon - \log(\bar{\Omega}_0) + \mathbb{E}_A [\log(\Omega(A))]}{\left(\mathbb{E}_A [\Delta(A)] - \log(\bar{\Omega}_0) \right)}.\quad (2.5.12)$$

Hence, our suboptimal fusion rule is simply the thresholding of the received signal, which is the same as the optimal rule in the case of the MF local detector above. The performance is again attained via Theorems 2.1 and 2.2 . Therefore

$$P_{FA} = \sum_{m=\Upsilon_2}^{\infty} \frac{\bar{\Omega}_0^m}{m!} e^{-\bar{\Omega}_0} \quad (2.5.13)$$

$$P_D = \sum_{m=\Upsilon_2}^{\infty} \mathbb{E}_A \left[\frac{\Delta^m(A)}{m!} e^{-\Delta(A)} \right] \quad (2.5.14)$$

which is exactly the same as the CR results.

2.5.2 Gaussian Channel

The next step is to introduce AWGN into the shared MAC channel. For simplicity, we set the SN's transmission power to unity. The received signal at the FC would be a corrupted version of that of (2.5.1), i.e.,

2.5. DECISION FUSION OVER SHARED MAC

$$\begin{aligned} Y &= \sum_{\mathbf{X}_i \in \Phi} I(\mathbf{X}_i) + V \\ &= \Omega + V \end{aligned} \quad (2.5.15)$$

where, as stated earlier, V is AWGN with zero mean and variance σ_c^2 . The LLR of a given received signal y is

$$\Lambda_{\text{G-ED}}(y) = \log \left(\frac{\mathbb{E}_1 \left[\exp \left(-\frac{1}{2\sigma_c^2} (y - \Omega)^2 \right) \right]}{\mathbb{E}_0 \left[\exp \left(-\frac{1}{2\sigma_c^2} (y - \Omega)^2 \right) \right]} \right) \quad (2.5.16)$$

where $\mathbb{E}_j[\cdot]$ is the expectation w.r.t Ω 's distribution under the \mathcal{H}_j hypothesis for $j = 0, 1$. Note that Ω 's distribution differs under \mathcal{H}_0 and \mathcal{H}_1 as Theorems 2.1 and 2.2 state. Obviously, it is not easy to analyze the distribution of $\Lambda_{\text{G-ED}}$ in (2.5.16). To obtain a suboptimal fusion rule, we proceed as before by applying Jensen's inequality to (2.5.16) leading to

$$\Lambda_{\text{G-ED}}(y) \leq \tilde{\Lambda}_{\text{G-ED}}(y) \quad (2.5.17)$$

$$= \mathbb{E}_1 \left[-\frac{1}{2\sigma_c^2} (y - \Omega)^2 \right] - \mathbb{E}_0 \left[-\frac{1}{2\sigma_c^2} (y - \Omega)^2 \right]. \quad (2.5.18)$$

Taking the upper bound above to be our fusion rule, it can be shown that

$$\tilde{\Lambda}_{\text{G-ED}}(y) = \frac{y}{\sigma_c^2} (\mathbb{E}_1[\Omega] - \mathbb{E}_0[\Omega]) - \frac{1}{2\sigma_c^2} (\mathbb{E}_1[\Omega^2] - \mathbb{E}_0[\Omega^2]) \quad (2.5.19)$$

which can be simply expressed as

$$\tilde{\Lambda}_{\text{G-ED}}(y) = y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \Upsilon_3 \quad (2.5.20)$$

where

$$\Upsilon_3 = \frac{\Upsilon - \frac{1}{2\sigma_c^2} (\mathbb{E}_1[\Omega^2] - \mathbb{E}_0[\Omega^2])}{\frac{1}{\sigma_c^2} (\mathbb{E}_1[\Omega] - \mathbb{E}_0[\Omega])}. \quad (2.5.21)$$

The communication noise changes the performance analysis when compared to (2.5.3) and (2.5.11). The distribution of Y in (2.5.15) is the convolution of the distributions of Ω and V , since it is the addition of the previous independent RVs. Thus it can be written as

2.5. DECISION FUSION OVER SHARED MAC

$$p_j^{\text{G-ED}}(y) = \frac{1}{\sqrt{2\pi\sigma_c^2}} \mathbb{E}_j \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega)^2 \right) \right] \quad (2.5.22)$$

where the expectation is taken w.r.t to (2.3.3) for $j = 0$ and w.r.t (2.3.5) for $j = 1$. Another layer of uncertainty is added to the system due to having communication noise. Hence, the global false alarms are due to two factors now: false local detection and communication noise. The performance can be improved of course by increasing the transmission power, which implies decreasing σ_c^2 . Asymptotically, as σ_c^2 decreases, the Gaussian distribution of the noise resembles a Dirac function and consequently the result from the distribution convolution will look more like the Ω 's distribution. Hence, we expect that the performance of the fusion rule (2.5.11) bounds that of (2.5.20). The global false alarm and detection probabilities follow as

$$P_{FA} = \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{\gamma_3}^{\infty} \mathbb{E}_0 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega)^2 \right) \right] dy \quad (2.5.23)$$

$$P_D = \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{\gamma_3}^{\infty} \mathbb{E}_1 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega)^2 \right) \right] dy. \quad (2.5.24)$$

2.5.3 Rayleigh Fading Channel

Here we consider the case of having communication noise and fading, which causes fluctuation in the magnitude and phase. The received signal then takes the form

$$\begin{aligned} Y_{RX} &= \sum_{\mathbf{X}_i \in \Phi_d} H(\mathbf{X}_i) + V_c \\ &= \Omega_R + V_c \end{aligned} \quad (2.5.25)$$

where $H(\mathbf{X}_i)$ is the complex channel coefficient between the SN located at \mathbf{X}_i and the FC that follows the circular Gaussian distribution $\mathcal{CN}(0, \sigma_h^2)$ and V_c is a circular AWGN noise, i.e., $\mathcal{CN}(0, \sigma_c^2)$. Note that the information carrying part of (2.5.25) can be written as

$$\Omega_R = \sum_{\mathbf{X}_i \in \Phi_d} (|H(\mathbf{X}_i)| \cos(\Psi(\mathbf{X}_i)) + j |H(\mathbf{X}_i)| \sin(\Psi(\mathbf{X}_i))) \quad (2.5.26)$$

2.5. DECISION FUSION OVER SHARED MAC

where $|H(\mathbf{X}_i)|$ is a Rayleigh RV and $\Psi(\mathbf{X}_i)$ is a uniform RV in the interval $[0, 2\pi]$. Note that Ω_R is the distorted version of Ω due to having phase noise, which results in destructive combining of the transmitted signals. However, since the imaginary part of the above carries redundant information, the FC takes only the real part, i.e.,

$$Y = \Re \{Y_{RX}\} \quad (2.5.27)$$

where $\Re \{\cdot\}$ is an operator that extracts the real part of the argument. The LLR in this instance is

$$\Lambda_{R-ED}(y) = \log \left(\frac{\mathbb{E}_1 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_R)^2 \right) \right]}{\mathbb{E}_0 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_R)^2 \right) \right]} \right). \quad (2.5.28)$$

The suboptimal fusion rule for the Rayleigh fading case is by the same procedure adopted before given by:

$$\tilde{\Lambda}_{R-ED}(y) = y \underset{\mathcal{H}_0}{\overset{\mathcal{H}_1}{\gtrless}} \Upsilon_4 \quad (2.5.29)$$

where Υ_4 takes a form similar to Υ_3 . To analyze the performance we observe that the received signal distribution in (2.5.25) is the convolution of Ω_R 's distribution and the Gaussian noise distribution. We can see that Ω_R is the random sum of the channel gains $H(\mathbf{X}_i)$, which are zero mean Gaussian RVs. As a result, the fusion rule's distribution in (2.5.29) given Ω (which is the number of the detecting SNs) is exactly Gaussian. Consequently, the received signal's distribution is the following Gaussian mixture

$$p_j^{R-ED}(y) = \mathbb{E}_j \left[\frac{1}{\sqrt{2\pi (\Omega\sigma_h^2 + \sigma_c^2)}} \exp \left(-\frac{y^2}{2\Omega\sigma_h^2 + 2\sigma_c^2} \right) \right]. \quad (2.5.30)$$

Note that the variance is a RV that depends on the number (Ω) of detecting SN. We expect the performance here to be worse than (2.5.20) because of the destructive combining in (2.5.25) due to the presence of phase noise.

The P_{FA} and P_D are consequently

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

$$P_{FA} = \int_{\mathbf{r}_4}^{\infty} \mathbb{E}_0 \left[\frac{1}{\sqrt{2\pi(\Omega\sigma_h^2 + \sigma_c^2)}} \exp\left(-\frac{y^2}{2\Omega\sigma_h^2 + 2\sigma_c^2}\right) \right] dy \quad (2.5.31)$$

$$P_D = \int_{\mathbf{r}_4}^{\infty} \mathbb{E}_1 \left[\frac{1}{\sqrt{2\pi(\Omega\sigma_h^2 + \sigma_c^2)}} \exp\left(-\frac{y^2}{2\Omega\sigma_h^2 + 2\sigma_c^2}\right) \right] dy. \quad (2.5.32)$$

As a conclusion to this section we note that the above fusion rules are fairly simple, intuitive, and work for any range of communication SNR in contrast to [89] and [39], which adopts different fusion rules for communication SNR.

2.6 Diversity Combining Techniques for MAC

In this section, we revise the centralized diversity combining scheme proposed in [82], in which the diversity processing is done at the FC. Then we introduce distributed diversity combining. In the latter, we first find the optimal distributed diversity combining scheme, which is unfortunately difficult to implement in practice. Hence, we propose two suboptimal methods, the distributed maximum ratio combining (dMRC) and the distributed equal gain combining (dEGC).

2.6.1 Centralized Diversity Combining

The conventional strategy to combat fading is to use diversity combining. However, diversity combining will not help if the received signal in (2.5.25), since the received signal has been destroyed by phase noise. Intuitively, the ED can be used in this case to extract information from the variance [76, 84]. Now, several instances of the MAC can be centrally combined at the FC [82]. The fusion rule therefore is

$$\Lambda_{\text{CDC}} = \sum_{b=0}^{B-1} |y_{RX}|^2 \quad (2.6.1)$$

where y_{RX} is a realization of the received signal in (2.5.25). In order to simplify the analysis of the above fusion rule, it can be written as

$$\Lambda_{\text{CDC}} = (\Omega\sigma_h^2 + \sigma_c^2) Z \quad (2.6.2)$$

where

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

$$Z = \sum_{b=0}^{B-1} \frac{|Y_b|^2}{(\Omega\sigma_h^2 + \sigma_c^2)} \quad (2.6.3)$$

is a central Chi square RV with B degrees of freedom, i.e., $Z \sim \chi_B^2$. Then, the cumulative distribution function (cdf) for the j th hypothesis is

$$\begin{aligned} P_j^{\text{CDC}}(y) &= \mathbb{P}(\Lambda_{\text{CDC}} \leq y; \mathcal{H}_j) \\ &= \mathbb{E}_j \left[\mathbb{P} \left(Z \leq \frac{y}{(\Omega\sigma_h^2 + \sigma_c^2)} \right) \right] \\ &= \mathbb{E}_j \left[P_Z \left(\frac{y}{(\Omega\sigma_h^2 + \sigma_c^2)} \right) \right] \end{aligned} \quad (2.6.4)$$

where P_Z is the cdf of χ_B^2 . Clearly, the distribution is complicated. Therefore, we suggest approximation by a Gamma distribution via moment matching to find the scale and shape parameters ($\alpha_{\text{CDC},j}$ and $\beta_{\text{CDC},j}$ respectively) of the Gamma distribution. The first and second moments can be shown to be

$$\mathbb{E}_j[\Lambda_{\text{CDC}}] = (\mathbb{E}_j[\Theta] \sigma_h^2 + \sigma_c^2) B \quad (2.6.5)$$

$$\text{var}_j(\Lambda_{\text{CDC}}) = 2B (\mathbb{E}_j[\Theta^2] \sigma_h^4 + 2\mathbb{E}_j[\Theta] \sigma_h^2 \sigma_c^2 + \sigma_c^4) \quad (2.6.6)$$

Consequently, we can approximate Λ_{CDC} by $\tilde{\Lambda}_{\text{CDC}}$ which has a Gamma distribution, i.e., $\tilde{\Lambda}_{\text{CDC}} \sim \text{Gamma}(\alpha_{\text{CDC},j}, \beta_{\text{CDC},j})$ for \mathcal{H}_j , $j = 0, 1$. Obviously, having higher diversity improves performance as can be conjectured from (2.6.5). However, increasing the diversity also increases the hardware complexity and size of the FC. Then the P_{FA} and P_D are:

$$P_{FA} \approx \mathbb{P}(\tilde{\Lambda}_{\text{CDC}} \geq \Upsilon; \mathcal{H}_0) \quad (2.6.7)$$

$$P_D \approx \mathbb{P}(\tilde{\Lambda}_{\text{CDC}} \geq \Upsilon; \mathcal{H}_1). \quad (2.6.8)$$

2.6.2 Distributed Diversity Combining

Considering the above disadvantages of conventional diversity combining, new ways of improving the performance should be pursued. In CDC, little can be done to process the received signal at the FC to mitigate the effects of fading, due to the aggregation of the transmitted signal in the shared MAC. In such a situation, the

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

FC cannot resolve the signals to perform diversity combining. Therefore, inspired by precoding we propose distributed diversity combining where the SN scales its transmitted signal by $U(\mathbf{X}_i)$ that is a function of the corresponding channel coefficient $H(\mathbf{X}_i)$ between the SN and the FC. The transmitted signals are combined by virtue of the MAC, in order that the received signal at the FC in this case is

$$Y_{RX} = \sum_{\mathbf{X}_i \in \Phi_d} H(\mathbf{X}_i)U(\mathbf{X}_i) + V_c. \quad (2.6.9)$$

As before, we are interested in the real part of the above, $Y = \Re\{Y_{RX}\}$. We wish to design the precoding factors $U(\mathbf{X}_i)$ in order to improve the overall detection performance. To this end, an appropriate optimality measure should be chosen and then find the set of $U(\mathbf{X}_i)$ values that optimizes it. We adopt the J-divergence of the received signal distributions under \mathcal{H}_1 and \mathcal{H}_0 . It measures the discrepancy between two distributions, hence it can be used as a detection quality indicator.

2.6.2.1 Optimal Distributed Diversity Combining

We investigate the problem of optimally choosing $U(\mathbf{X}_i)$ in order to maximize the J-divergence, which is defined as

$$J(p_0(y), p_1(y)) = \mathbb{E}_1 \left[\log \left(\frac{p_1(y)}{p_0(y)} \right) \right] - \mathbb{E}_0 \left[\log \left(\frac{p_1(y)}{p_0(y)} \right) \right]. \quad (2.6.10)$$

Now we need to find a relationship between the J-divergence and the precoding factor $U(\mathbf{X}_i)$'s. However, this problem in the current form is difficult to solve since the J-divergence has a rather complicated form. Therefore, we resort to approximate the J-divergence by its upper bound (\tilde{J}), since it is known to be a tight bound. The following lemma derives the required upper bound.

Lemma 2.2. *Let the received signal at the FC take the form of*

$$Y = \Omega_f + W \quad (2.6.11)$$

where W is zero mean AWGN with variance σ^2 and Ω_f is defined for any positive function $f(\cdot)$ as

$$\Omega_f = \sum_{\mathbf{X}_i \in \Phi_d} f(\mathbf{X}_i). \quad (2.6.12)$$

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

Then the J-divergence of the received signal is upper bounded by

$$J(p_0(\mathbf{y}), p_1(\mathbf{y})) \leq \frac{1}{2\sigma^2} (\mathbb{E}_1[\Omega_f] - \mathbb{E}_0[\Omega_f])^2. \quad (2.6.13)$$

Proof. See Appendix A.6. □

Recall that the expectations in the LHS of (2.6.13) are defined as

$$\begin{aligned} \mathbb{E}_0[\Omega_f] &= \mathbb{E}_0 \left[\sum_{\mathbf{x}_i \in \Phi_d} f(\mathbf{X}_i) \right] \\ &= \lambda P_{fa} \int_{\mathcal{F}} f(\mathbf{x}) d\mathbf{x} \end{aligned} \quad (2.6.14)$$

and

$$\mathbb{E}_1[\Omega_f] = \lambda \int_{\mathcal{F}} f(\mathbf{x}) \bar{P}_d(\mathbf{x}) d\mathbf{x} \quad (2.6.15)$$

where

$$\bar{P}_d(\mathbf{x}) = \mathbb{E}_A [P_d(\mathbf{x}, A)] \quad (2.6.16)$$

is defined for convenience.

From (2.6.9) we can set $f(\mathbf{X}_i) = H(\mathbf{X}_i)U(\mathbf{X}_i)$ and then find the J-divergence's upper bound. However, since the CSI is assumed to be known at the i th SN, $f(\mathbf{x}_i) = h(\mathbf{x}_i)u(\mathbf{x}_i)$. Then the J-divergence upper bound is

$$\tilde{J}(u) = \frac{\lambda^2}{\sigma_c^2} \left(\int_{\mathcal{F}} u(\mathbf{x})h(\mathbf{x})\bar{P}(\mathbf{x})d\mathbf{x} \right)^2 \quad (2.6.17)$$

where

$$\bar{P}(\mathbf{x}) = \bar{P}_d(\mathbf{x}) - P_{fa}. \quad (2.6.18)$$

Having such an expression that connects the precoding factor (u) with \tilde{J} , we now provide the optimal $u(\cdot)$ that maximizes $\tilde{J}(u)$ in the following lemma.

Lemma 2.3. *The solution of*

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

$$\hat{u}(\mathbf{x}) = \arg \max_{u(\mathbf{x})} \bar{J}(u) \quad (2.6.19)$$

is

$$\hat{u}(\mathbf{x}) = kh^*(\mathbf{x})\bar{P}(\mathbf{x}) \quad (2.6.20)$$

where $h^*(x)$ is the conjugate of the channel coefficient and (k) is an arbitrary constant.

Proof. If no constraints are applied, then using the Cauchy-Schwarz inequality directly yields (2.6.20). \square

The above result implies that optimal pre-equalization means more energy should be pumped into the SNs with good channel and good local detection. Another interesting feature of (2.6.20) is that we have a degree of freedom provided by k . We can set it to

$$k = k_0 \frac{\sigma_c}{\lambda} \quad (2.6.21)$$

where k_0 is an arbitrary constant as well. However, the above choice is intuitive from two aspects. First, the transmitted signal's power should be proportional to the channel noise in order to overcome it. And second, the MAC aggregation (if the phase noise is compensated) leads to a beamforming effect. Therefore, the desired SNR at FC is divided between the SNs. This is reflected in (2.6.21) by dividing by λ .

2.6.2.2 Distributed Maximum Ratio Combining (dMRC)

Finding $\bar{P}(\mathbf{x})$ nonetheless requires the knowledge of the target parameters, which is unattainable in practice. Thus, we propose the suboptimal pre-equalization at a given SN located at \mathbf{X}_i

$$u_{\text{dMRC}}(\mathbf{X}_i) = kh^*(\mathbf{X}_i). \quad (2.6.22)$$

Note that we are using small letters notation, which implies that the variables are deterministic. This is due to assuming the knowledge of the CSI at all the detecting SNs.

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

The above indeed resembles the maximum ratio combining (MRC) technique. However, in our case the the channel is *distributed* among the SNs that pre-scale the transmitted signal by (2.6.22), hence the name. Let the number of detecting SNs be M , and let for convenience the channel coefficient and precoding factor for the SN located at \mathbf{x}_i be denoted as $h(i)$ and $u_{\text{dMRC}}(i)$ respectively, then the received signal at the FC (the real part of it):

$$\begin{aligned}
 Y &= \Re \left\{ \sum_{i=0}^{M-1} u_{\text{dMRC}}(i)h(i) + V_c \right\} \\
 &= \Re \left\{ \sum_{i=0}^{M-1} kh^*(i)h(i) + V_c \right\} \\
 &= k \sum_{i=0}^{M-1} |h(i)|^2 + V_r
 \end{aligned} \tag{2.6.23}$$

where V_r is the real component of V_c . To analyze the performance, we consider the stochastic geometric framework of the network, i.e., we replace the summation in (2.6.23) with the summation over the network realization. In other words,

$$Y = \underbrace{\sum_{\mathbf{X}_i \in \Phi_d} |H(\mathbf{X}_i)|^2}_{\Omega_{\text{dMRC}}} + V_r \tag{2.6.24}$$

where $k = 1$ for simplicity. Y here is the summation of a Gaussian RV and Ω_{dMRC} which is a RV resulting from a random sum. It is well known that, if $|H(\mathbf{X}_i)|$ is Rayleigh with power of σ_h^2 , then the sum of say N replicas is a Gamma RV with shape and scale factor of M and $2\sigma_h^2$ respectively. In other words, the distribution of Ω_{dMRC} (given M) is $\Gamma(M, 2\sigma_h^2)$. Consequently, Ω_{dMRC} is a Gamma mixture, with mixing coefficients being the probabilities of the number of detecting SNs. So, the distribution of the received signal is

$$p_j^{\text{dMRC}}(y) = \frac{1}{\sqrt{2\pi\sigma_c^2}} \mathbb{E}_j \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_{\text{dMRC}})^2 \right) \right]. \tag{2.6.25}$$

Consequently, the P_{FA} and P_D are exactly

2.6. DIVERSITY COMBINING TECHNIQUES FOR MAC

$$P_{FA} = \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{\mathbf{r}}^{\infty} \mathbb{E}_0 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_{\text{dMRC}})^2 \right) \right] dy \quad (2.6.26)$$

$$P_D = \frac{1}{\sqrt{2\pi\sigma_c^2}} \int_{\mathbf{r}}^{\infty} \mathbb{E}_1 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_{\text{dMRC}})^2 \right) \right] dy. \quad (2.6.27)$$

2.6.2.3 Distributed Equal Gain Combining (dEGC)

Equation (2.6.22) is nothing but the MRC distributed amongst the SN. However, (2.6.22), as in (2.6.20), requires the full knowledge of the channel. Relaxing this requirement to knowing only the channel phase yields the distributed EGC

$$u_{\text{dEGC}}(\mathbf{X}_i) = k \exp(-j \arg \{h(\mathbf{X}_i)\}) \quad (2.6.28)$$

where $\arg \{\cdot\}$ is the argument of the complex channel.

Following the same methodology in dMRC above to analyze the performance, the real part of the received signal is

$$Y = \underbrace{\sum_{\mathbf{X}_i \in \Phi_d} |H(\mathbf{X}_i)|}_{\Omega_{\text{dEGC}}} + V_r \quad (2.6.29)$$

where Ω_{dEGC} is a random sum of Rayleigh RVs. However, there is no closed form expression for the distribution of detecting SNs. Thus, we propose approximating Ω_{dEGC} by a Gamma distribution as was done earlier. The first and second moments are

$$\begin{aligned} \mathbb{E}_j[\Omega_{\text{dEGC}}] &= \mathbb{E}_j \left[\sum_{\mathbf{X}_i \in \Phi_d} |H(\mathbf{X}_i)| \right] \\ &= \lambda \int_{\mathcal{F}} \int_0^{\infty} |h(\mathbf{x})| p_H(h) \bar{P}_d(\mathbf{x}) dh d\mathbf{x} \\ &= \left(\int_0^{\infty} |h| p_H(h) dh \right) \left(\lambda \int_{\mathcal{F}} \bar{P}_d(\mathbf{x}) d\mathbf{x} \right) \end{aligned} \quad (2.6.30)$$

$$= \mathbb{E}_j[\Omega] \sqrt{\frac{\pi\sigma_h^2}{2}}. \quad (2.6.31)$$

In (2.6.30), the integration is decoupled because $h(\mathbf{X}_i)$ is independent of \mathbf{X}_i . In a similar manner, the second moment is

2.7. SIMULATION RESULTS AND DISCUSSIONS

$$\text{var}_j [\Omega_{\text{dEGC}}] = 2\sigma_h^2 \mathbb{E}_j [\Omega]. \quad (2.6.32)$$

After computing the scale and shape factors $\alpha_{\text{dEGC},j}$ and $\beta_{\text{dEGC},j}$, now Ω_{dEGC} 's distribution is approximated by Gamma ($\alpha_{\text{dEGC},j}, \beta_{\text{dEGC},j}$), where $\alpha_{\text{dEGC},j}$ and $\beta_{\text{dEGC},j}$ are the scale and shape factors respectively. Finally, the received signal distribution is

$$p_j^{\text{dEGC}}(y) \approx \frac{1}{\sqrt{2\pi\sigma_c^2}} \mathbb{E}_j \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_{\text{dEGC}})^2 \right) \right]. \quad (2.6.33)$$

The P_{FA} and P_D are approximated by

$$P_{FA} \approx \int_{\mathcal{Y}} \mathbb{E}_0 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_{\text{dEGC}})^2 \right) \right] dy \quad (2.6.34)$$

$$P_D \approx \int_{\mathcal{Y}} \mathbb{E}_1 \left[\exp \left(-\frac{1}{\sigma_c^2} (y - \Omega_{\text{dEGC}})^2 \right) \right] dy. \quad (2.6.35)$$

It is worth noting that the distributions (2.5.22), (2.6.25), and (2.6.33) are the convolution of a Gaussian RV with a random sum RV. We can think of the Gaussian distribution as an interpolation operator on the random sum distribution, which tends to take a shape closer to a discrete distribution than a continuous one. Therefore, having a low SNR channel would in fact produce a ‘‘nicer’’ fusion rule distribution.

2.7 Simulation Results and Discussions

2.7.1 Optimal Fusion Rule

We show the validity of the theoretical expressions for the OFR moments derived in Subsection 2.4. We simulate a WSN in a field of $300 \times 300 m^2$ with SNs deployed according to a uniform random distribution therein. To exclude the edge effects we only choose SNs within $150 m$ from the origin. The SNs are deployed with intensity $\lambda = 10^{-2} \text{SN}/m$. Every SN takes 200 samples of a target located at the origin, without loss of generality. The SNR is defined to be the target's emitted signal power over the noise power at the target's location, i.e., it is a^2/σ^2 . We run the simulation 2×10^5 Monte Carlo iterations for SNRs of 10 and 20 dB.

2.7. SIMULATION RESULTS AND DISCUSSIONS

Figure 2.7.1 shows a perfect match between the theoretical and the simulated mean of OFR under both hypotheses and for SNRs of 10 and 20 dB. The mean increases as P_{fa} increases since more SNs are active then. For the same reason the mean under \mathcal{H}_1 is greater than \mathcal{H}_0 . In a similar manner, the theoretical and simulated variances fit closely in Figure 2.7.2. However, the variance decreases with increasing P_{fa} under \mathcal{H}_1 in contrast to the \mathcal{H}_0 case. The reason for the later is that the mean and variance are linearly proportional to P_{fa} . In contrast with the former case under \mathcal{H}_1 , because having greater P_{fa} leads to more detecting SNs, since the local threshold is higher, that in turn reduces the uncertainty about the target, and hence reduces the OFR variance.

Finally, Figures 2.7.3, and 2.7.4 depict the ROC of the OFR for $P_{fa} = 10^{-3}$ and 10^{-2} . Although the Gamma distribution provides better approximation, both distributions yield similar global detection performance. Increasing the SNR, on the other hand and the approximation is improved due to more SNs being involved.

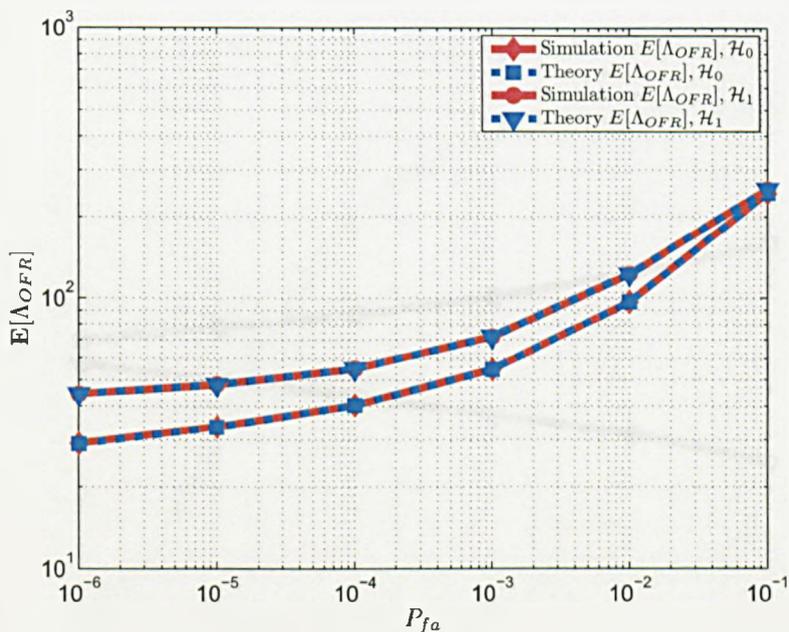
2.7.2 Fusion Rules over Fading Shared Channel

Now we show the theoretical and simulated performance of the CR, the fusion rules in Subsection 2.5, and the diversity combining techniques of subsection 2.6. We simulate a WSN in a field of $300 \times 300 m^2$ with SNs deployed randomly there and every SNs takes 200 samples like before. In the latter we set the target's amplitude distribution according to a Gaussian distribution with different mean and variance values. The observation SNR is defined to be the target's emitted signal power over the noise power at the target's location, i.e., it is $\mathbb{E}[A^2]/\sigma_s^2$, which is chosen to be 18 dB. The Rayleigh channel on the other hand, has unity power, i.e., $\sigma_h^2 = 1$. The transmitted SNR_c is defined to be $1/\sigma_c^2$. The simulation is run for 10^4 Monte Carlo iterations.

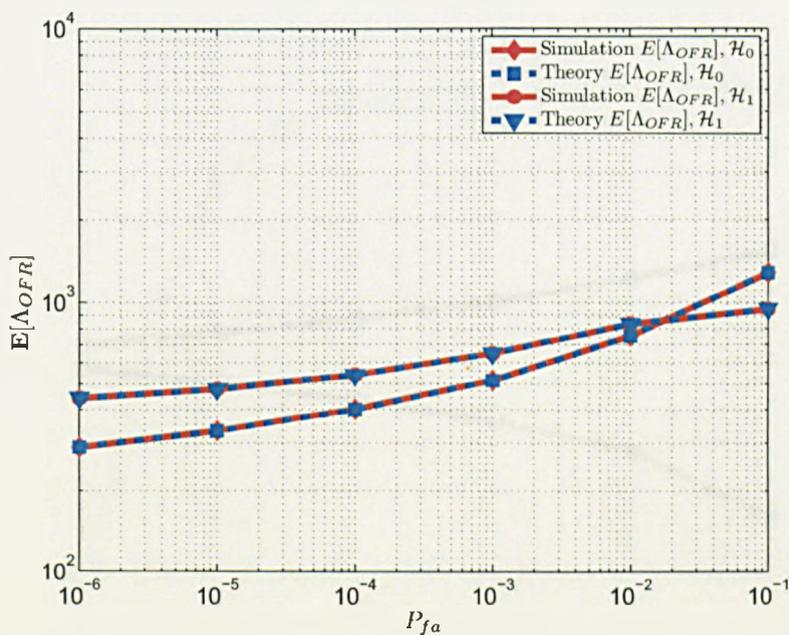
In Fig. 2.7.5, the ROC curves are shown for all the discussed scenarios for a target amplitude following $\mathcal{N}(20, 8)$ and sensing SNR of 20 and 25dB. Whereas Fig 2.7.6 shows the ROC graphs for the case of target amplitude distribution $\mathcal{N}(10, 5)$ and sensing SNR of 20 and 15dB. The predicted theoretical performance closely matches the simulation. Even in the case of dEGC and CDC, which are based on approximation, the theoretical expressions are very close to the simulation.

The best performance, obviously, is given by the use of the MF with ideal channels. In fact it is perfect due to the use of exact information about the target. Using the ED with ideal channels however, leads to some degradation in performance because statistical information is employed instead. Further degradation is caused in

2.7. SIMULATION RESULTS AND DISCUSSIONS



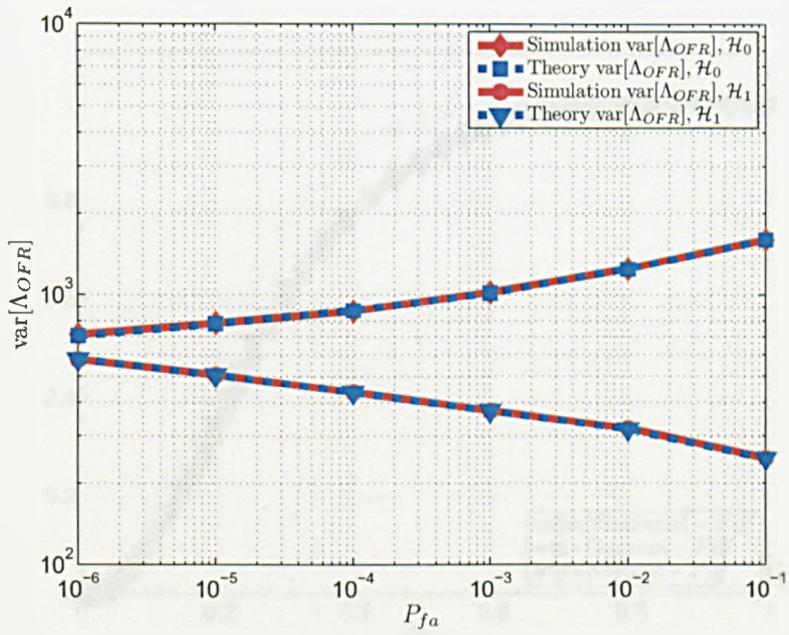
(a) $\text{SNR}_s = 10$ dB.



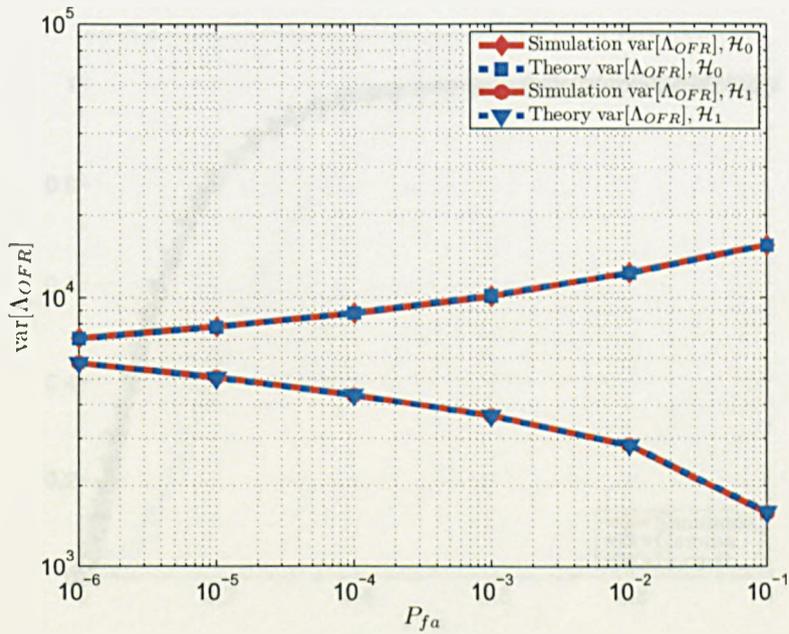
(b) $\text{SNR}_s = 20$ dB.

Figure 2.7.1: The mean of the optimal fusion rule (OFR) ($E[\Lambda_{\text{OFR}}]$) plotted against the local probability of false alarm, P_{fa} , under both null and alternative hypotheses, \mathcal{H}_0 and \mathcal{H}_1 respectively. The solid line represents simulated values and the dashed line is the theoretical values.

2.7. SIMULATION RESULTS AND DISCUSSIONS



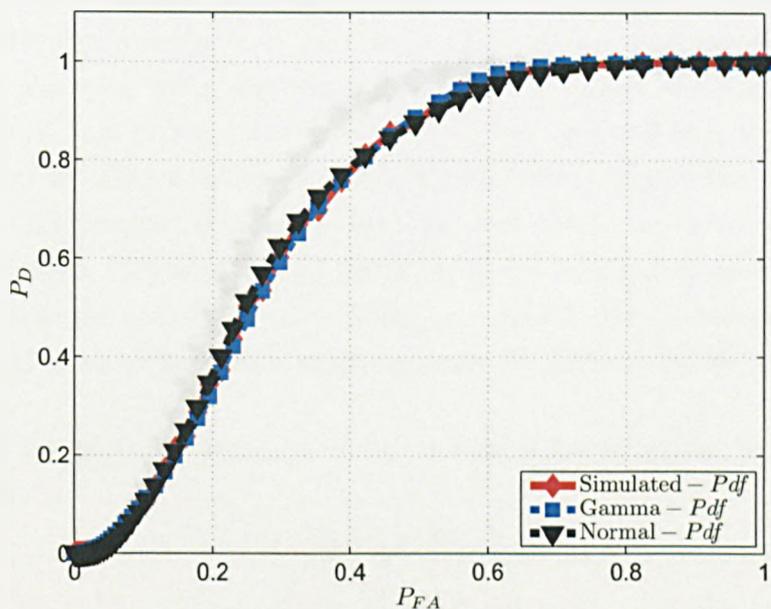
(a) $\text{SNR}_s = 10$ dB.



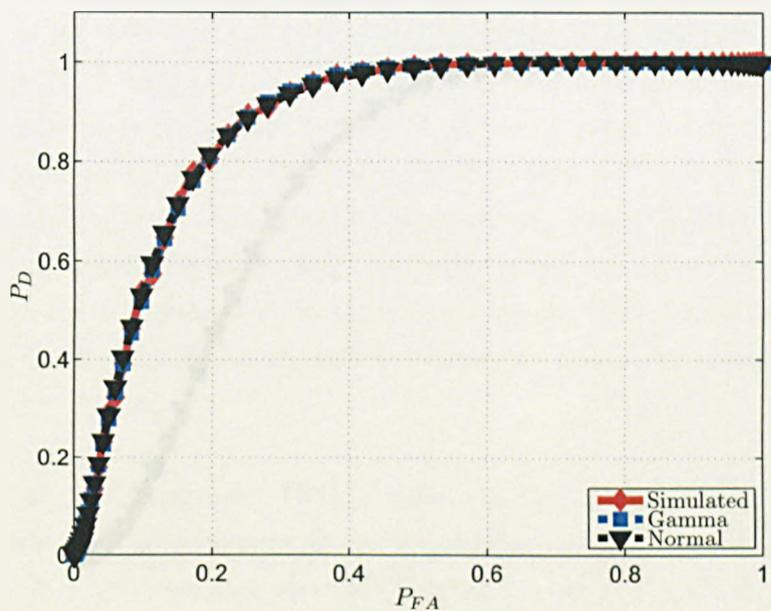
(b) $\text{SNR}_s = 20$ dB.

Figure 2.7.2: The variance of the optimal fusion rule (OFR) ($\mathbb{E}[\Lambda_{OFDR}]$) plotted against the local probability of false alarm, P_{fa} , under both null and alternative hypotheses, \mathcal{H}_0 and \mathcal{H}_1 respectively. The solid line represents simulated values and the dashed line is the theoretical values.

2.7. SIMULATION RESULTS AND DISCUSSIONS



(a) $\text{SNR}_s = 10$ dB.



(b) $\text{SNR}_s = 20$ dB.

Figure 2.7.3: ROC for optimal fusion rule (OFR) at $P_{fa} = 10^{-3}$. Solid line represents the Gamma distribution approximation and the dashed line represents the normal approximation.

2.7. SIMULATION RESULTS AND DISCUSSIONS

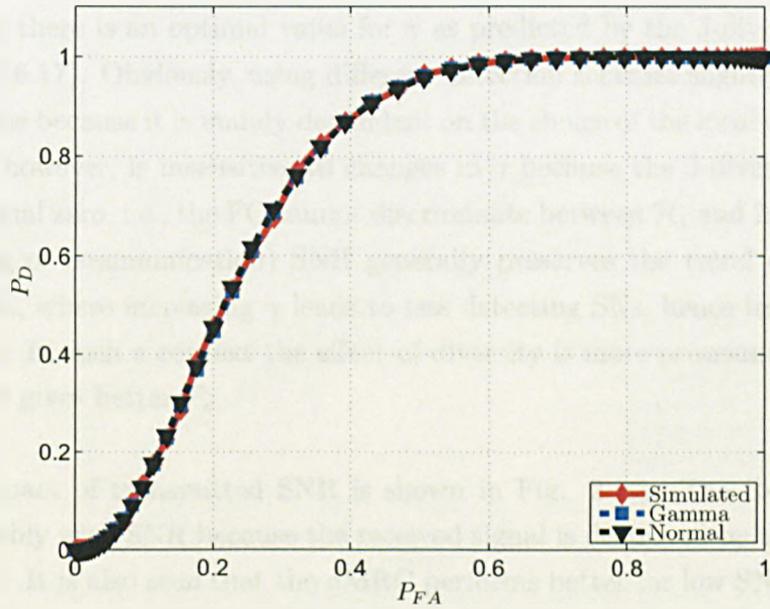
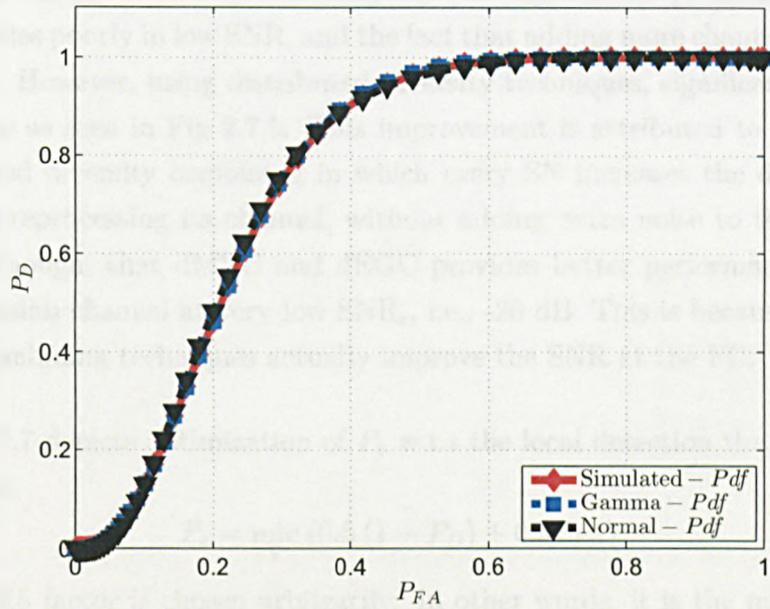


Figure 2.7.4: ROC for optimal fusion rule (OFR) at $P_{fa} = 10^{-2}$. Solid line represents the Gamma distribution approximation and the dashed line represents the normal approximation.

2.7. SIMULATION RESULTS AND DISCUSSIONS

the case of, Gaussian channels. On the other hand, Rayleigh fading channels yields the worse performance as the figures show in which we have $P_D = P_{FA}$. Using CDC with diversity order of 10 provides a slight improvement. This is mainly due to using the energy detector, which only uses the signal's energy for detection, and hence operates poorly in low SNR, and the fact that adding more channels introduces more noise. However, using distributed diversity techniques, significantly improves performance as seen in Fig 2.7.5. This improvement is attributed to the structure of distributed diversity combining in which every SN increases the diversity gain, by locally preprocessing its channel, without adding extra noise to the system. It is noticed though, that dMRC and dEGC provides better performance compared to the Gaussian channel at very low SNR_c, i.e., -20 dB. This is because distributed diversity combining techniques actually improve the SNR at the FC.

Fig. 2.7.7 depicts optimization of P_e w.r.t the local detection threshold, γ . We define P_e as

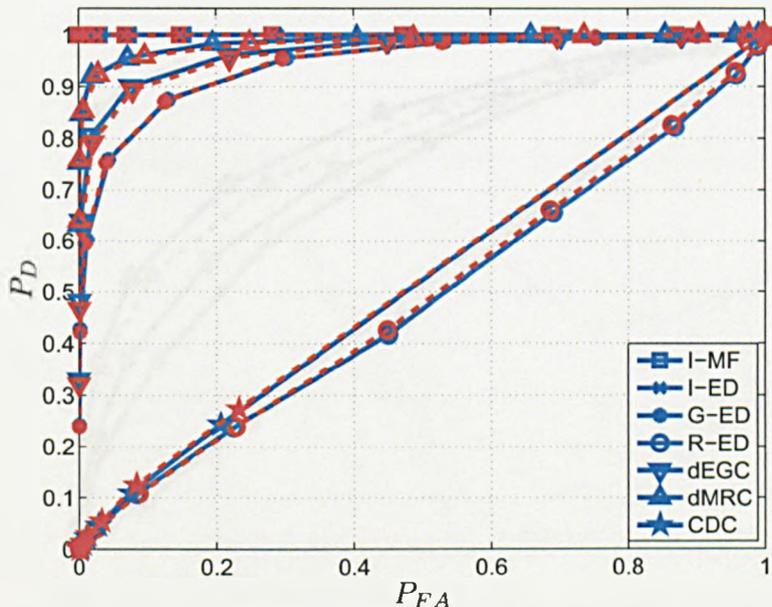
$$P_e = \min_{\gamma} (0.5 (1 - P_D) + 0.5 P_{FA}) \quad (2.7.1)$$

where the 0.5 factor is chosen arbitrarily. In other words, it is the minimum error probability over the global detection threshold.

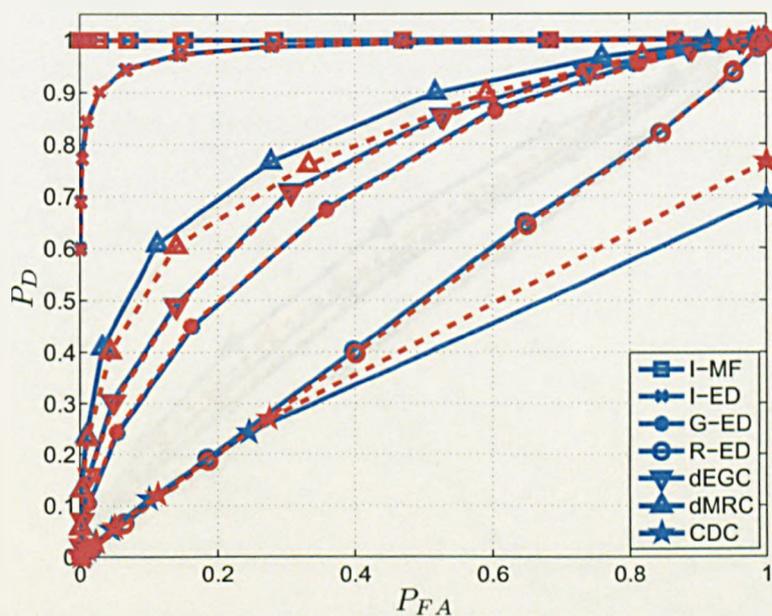
Clearly, there is an optimal value for γ as predicted by the J-divergence upper bound in (2.6.17). Obviously, using different detection schemes slightly changes the optimal value because it is mainly dependent on the choice of the local detector. The R-ED case however, is insensitive to changes in γ because the J-divergence can be shown to equal zero, i.e., the FC cannot discriminate between \mathcal{H}_1 and \mathcal{H}_0 . Increasing the (sensing or communication) SNR generally preserves the trend except in the case of CDC, where increasing γ leads to less detecting SNs, hence less destructive interference. In such a context the effect of diversity is more pronounced, therefore increasing γ gives better P_e .

The impact of transmitted SNR is shown in Fig. 2.7.8. The R-ED does not change notably with SNR because the received signal is destroyed by the incoherent interference. It is also seen that the dMRC performs better for low SNRs, while the dEGC outperforms it when the channel is more reliable. At higher SNRs the G-ED approaches I-ED. However, all the fusion rules level-off due to the channel effect. As predicted in Fig. 2.7.7, increasing γ provides better performance especially in the CDC case, which even outperforms I-ED.

2.7. SIMULATION RESULTS AND DISCUSSIONS



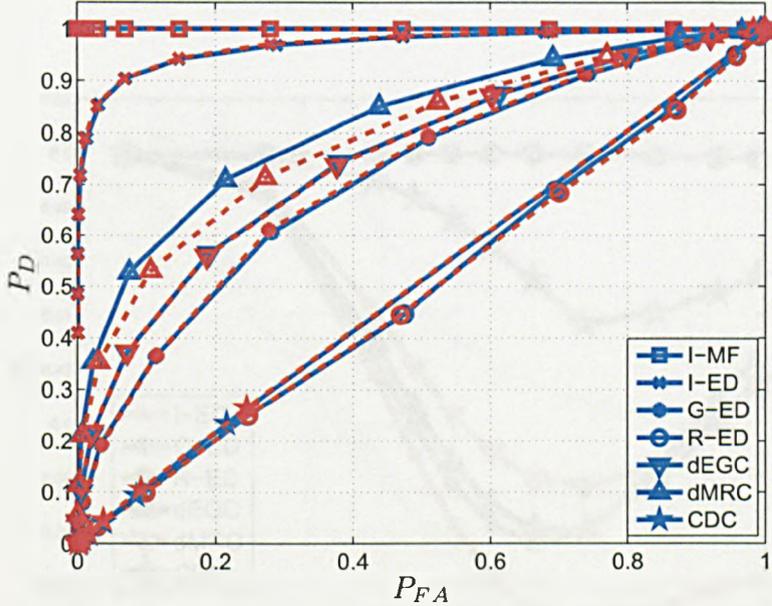
(a) The target amplitude follows the Gaussian distribution $\mathcal{N}(20, 8)$ and the sensing SNR (SNR_s) is 25dB.



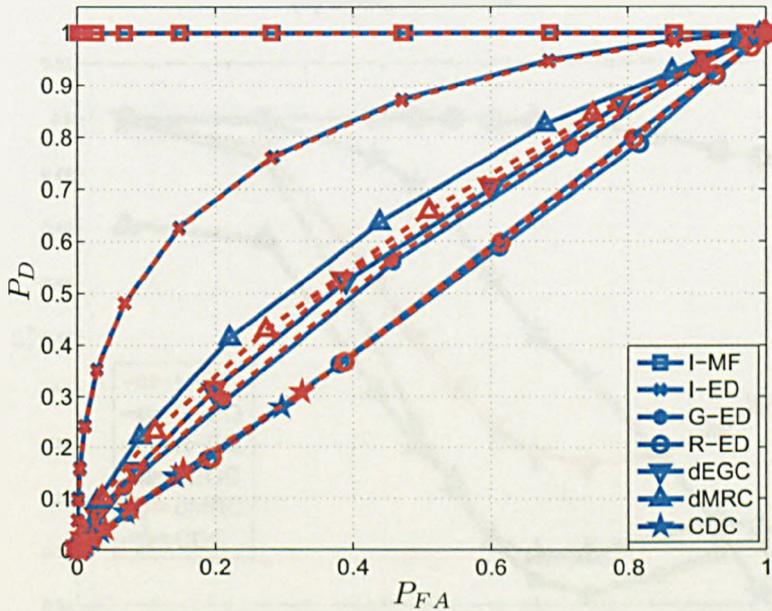
(b) The target amplitude follows the Gaussian distribution $\mathcal{N}(20, 8)$ and the sensing SNR (SNR_s) is 20dB.

Figure 2.7.5: Global probability of detection, P_D , against global probability of false alarm, P_{FA} . Solid line represents the theoretical values, and dashed line represents the simulation values. Network deployment density $\lambda = 5 \times 10^{-3}$, channel noise $\text{SNR}_c = -20\text{dB}$, Rayleigh channel power $\sigma_h^2 = 0\text{dB}$, and local detection threshold $\gamma = 235$ ($P_{fa} = 10^{-2}$).

2.7. SIMULATION RESULTS AND DISCUSSIONS



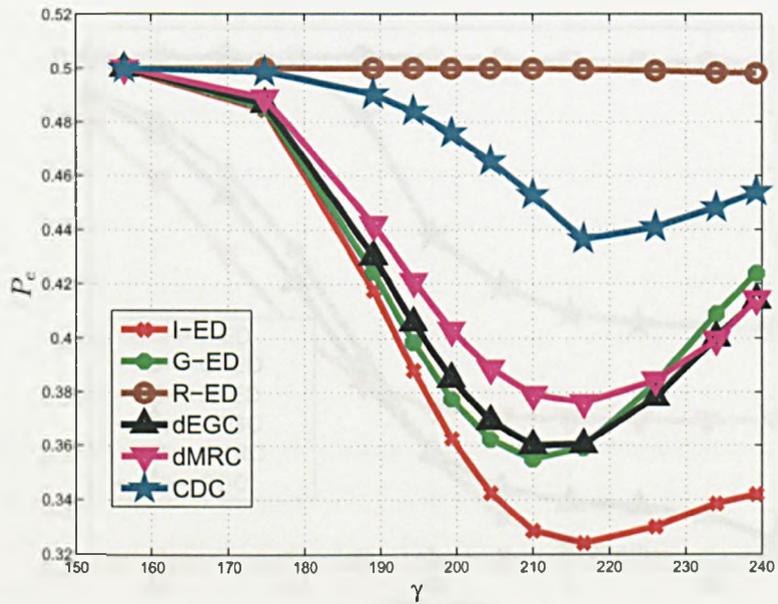
(a) The target amplitude follows the Gaussian distribution $\mathcal{N}(10, 5)$ and the sensing SNR (SNR_s) is 20dB.



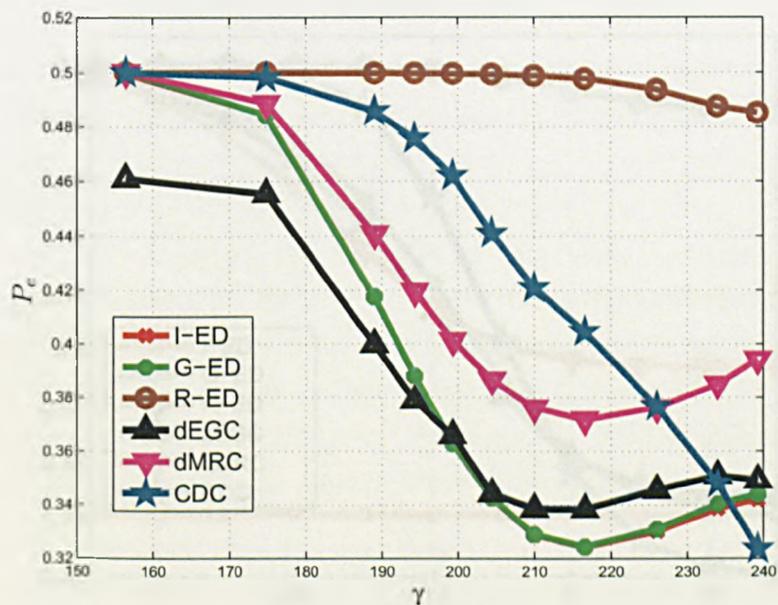
(b) The target amplitude follows the Gaussian distribution $\mathcal{N}(10, 5)$ and the sensing SNR (SNR_s) is 15dB.

Figure 2.7.6: Global probability of detection, P_D , against global probability of false alarm, P_{FA} . Solid line represents the theoretical values, and dashed line represents the simulation values. Network deployment density $\lambda = 5 \times 10^{-3}$, channel noise $\text{SNR}_c = -20\text{dB}$, Rayleigh channel power $\sigma_h^2 = 0\text{dB}$, and local detection threshold $\gamma = 235$.

2.7. SIMULATION RESULTS AND DISCUSSIONS



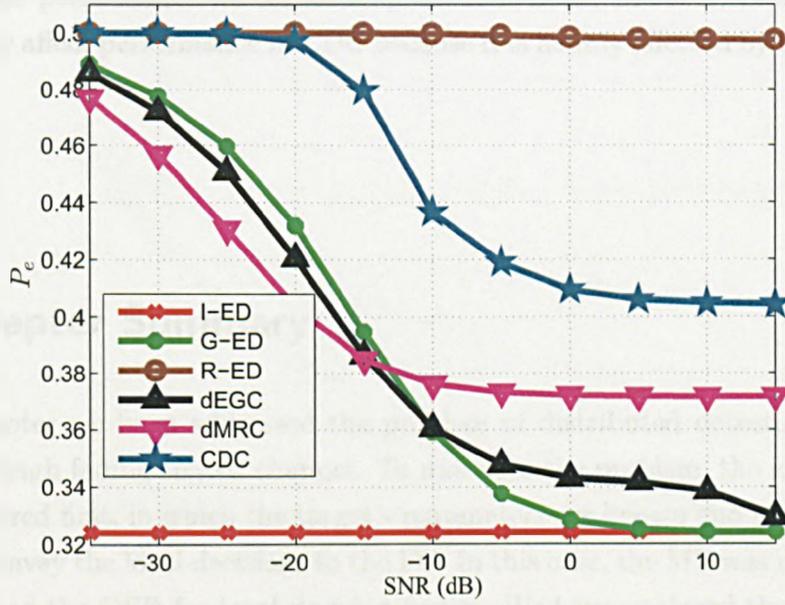
(a) $\text{SNR}_c = -10$ dB



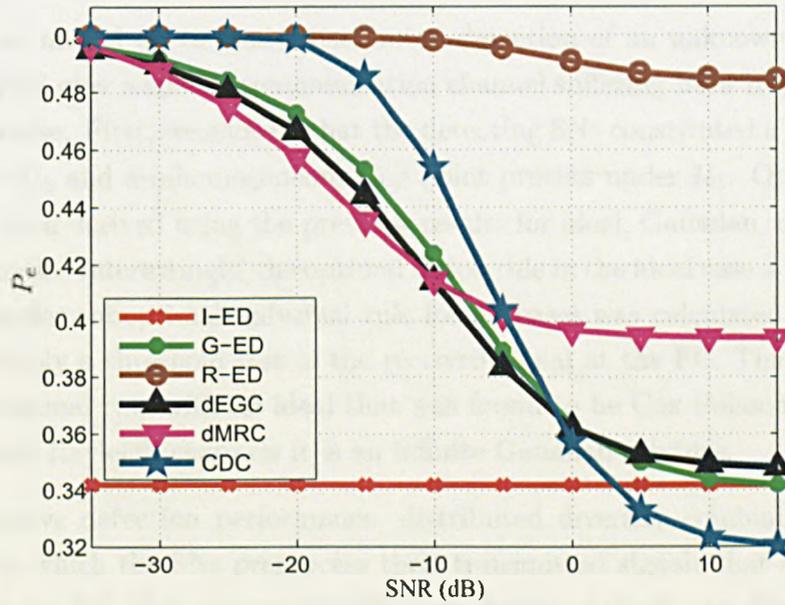
(b) $\text{SNR}_c = 10$ dB

Figure 2.7.7: Probability of error, P_e , against local detection threshold, γ , at $\sigma_h^2 = 1$ and $\lambda = 1 \times 10^{-3}$.

2.7. SIMULATION RESULTS AND DISCUSSIONS



(a) $\gamma = 216$



(b) $\gamma = 240$

Figure 2.7.8: Probability of error, P_e , against transmission SNR_c, at $\sigma_h^2 = 1$ and $\lambda = 1 \times 10^{-3}$.

2.8. CHAPTER SUMMARY

The effect of the SNs deployment density λ is depicted in Fig 2.7.9. As expected, more SNs results in better performance approaching perfect detection except in the Rayleigh fading channel case, again, because the received signal is destroyed by the random phase that causes destructive interference. Increasing the SNR clearly improves the performance for all techniques. It is noted however, that λ does not significantly affect performance in CDC because it is mainly affected by the channels' SNR.

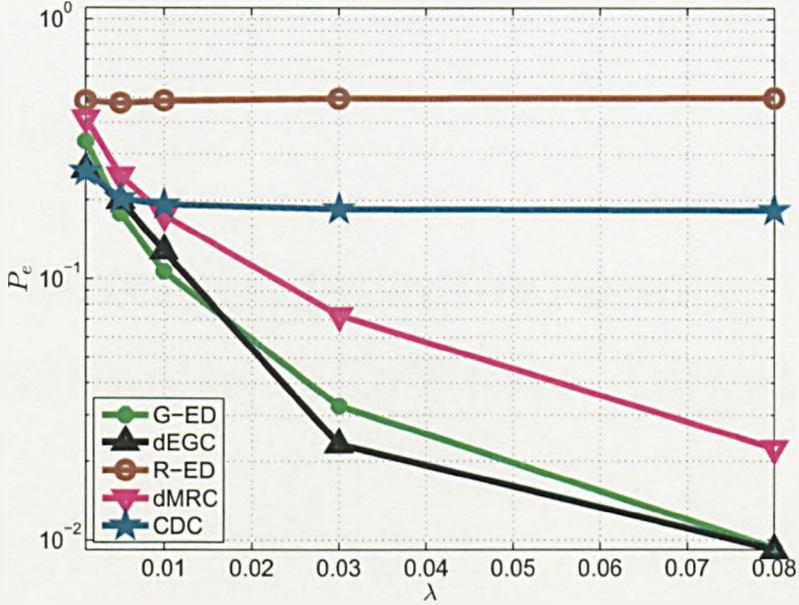
2.8 Chapter Summary

In this chapter we have addressed the problem of distributed detection in WSNs over a Rayleigh fading shared channel. To motivate the problem, the idealized case was considered first, in which the target's parameters are known and perfect parallel channels convey the local decisions to the FC. In this case, the MF was used for local detection and the OFR for local decision fusion. We have analyzed the statistics of the OFR for distributed detection using stochastic geometry.

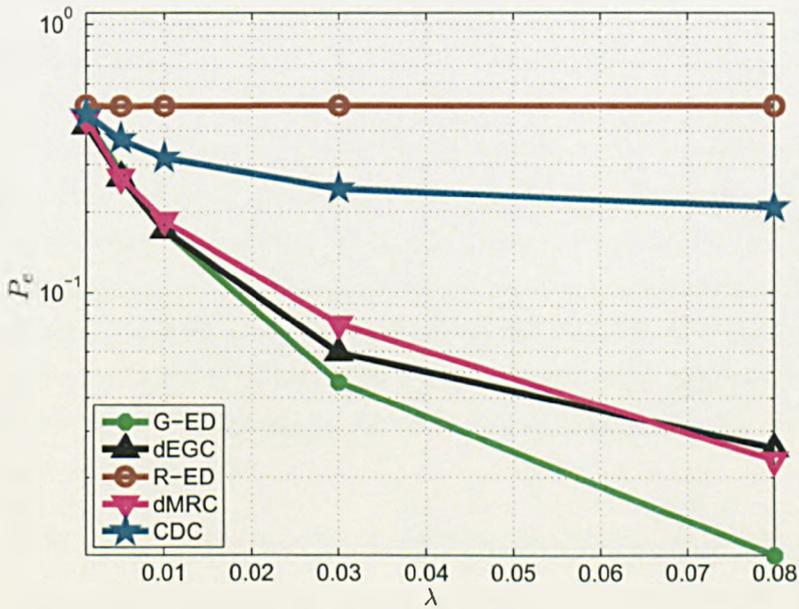
Then we moved on to study distributed detection of an unknown target with censored WSN over a shared communication channel suffering from Rayleigh fading and phase noise. First, we showed that the detecting SNs constituted a homogenous PPP under \mathcal{H}_0 and a inhomogeneous Cox point process under \mathcal{H}_1 . Optimal fusion rules have been derived using the previous results for ideal, Gaussian, and Rayleigh fading channels. Interestingly, the optimal fusion rule in the ideal case is the counting rule. A simpler suboptimal universal rule for all cases was calculated that turned up to be simply a threshold test of the received signal at the FC. The distribution of the suboptimal rule with an ideal that was found to be Cox Poisson, whereas in Gaussian and Rayleigh channels it is an infinite Gaussian mixture.

To improve detection performance, distributed diversity combining has been suggested in which the SNs preprocess their transmitted signals that are fused "in channel" at the FC. This scheme simplifies the design of the FC by distributing the processing among the SNs. In contrast to using CDC MAC, distributed diversity combining provides improved performance when increasing the number of SNs. We find the optimal technique by maximizing the J-divergence of the distribution of the received signal under the two hypotheses. The optimal technique turns out to be MRC weighted by the local decision confidence. Since the optimal fusion rule is difficult to find in practice, distributed MRC and EGC were proposed as suboptimal

2.8. CHAPTER SUMMARY



(a) $\text{SNR}_c = 10$



(b) $\text{SNR}_c = -10$

Figure 2.7.9: Probability of error, P_e , against network deployment density, λ , with $\sigma_h^2 = 1$, and $\gamma = 235$.

2.8. CHAPTER SUMMARY

alternatives. The performance of all fusion rules was exactly found, except in the case of dEGC, which was approximated. Simulations showed significant match between the theoretical and simulated results.

Chapter 3

Distributed Localization and Estimation of a Diffusive Target¹

3.1 Introduction

3.1.1 Motivation

WSNs are increasingly used in time-critical applications for surveillance and monitoring due to their low cost, scalability, and robustness. This class of applications requires a fast response to sudden events, e.g., enemy intrusion [19], natural disasters [13], or monitoring critical civilian structures [11]. A broad spectrum of phenomena can be described by diffusion [55]. In such scenarios, the diffusion measurements are used to localize any source or target and estimate relevant parameters. However, using WSNs for time-critical applications imposes several constraints on the system. The following are of interest:

- C1) Low bandwidth available at the SNs for wireless communication.
- C2) Low power available on board the SN.
- C3) Limited computational capability.
- C4) Timely target localization and estimation of its parameters.

¹Which is a target that generates a signal following the diffusion law, such as material concentration or heat.

3.1. INTRODUCTION

Moreover, the performance of such algorithms should be investigated under the previous operational constraints. The first three of them are the consequence of the nature of the SN, which is desired to be cheap and powered by on-board batteries. Hence, the communication and computational operations are restricted to affordable levels. The last constraint though, is dictated by the requirement of time-critical applications.

In this sequel we review the literature for related work aimed at solving the problem of distributed localization and estimation in WSNs and investigate its compliance with constraints C1-C4. Thus, the design of any localization and estimation algorithm for WSNs should take the above constraints into consideration.

3.1.2 Related Work

There exist several main directions of research in handling the problem of distributed localization and estimation of a diffusive source. All start from the partial differential equations (PDEs) governing the diffusion process then formulate it into a suitable model for analysis. The models relevant to this work are categorized as the following:

3.1.2.1 Static Parametric Model

In this model, the PDE is solved and the solution in turn is used to formulate a statistical parametric model of the measured data. One of the earliest works on this approach is due to Nehorai et. al. [90], in which the diffusion equation is solved for a stationary point source emitting vapor continuously under different boundary conditions. Detection and localization of the source is formulated according to GLRT and MLE respectively. The MLE results in a nonlinear and non-convex optimization problem that needs a good initialization in order to converge to the global minimum. In order to find a good initialization, the authors suggest waiting until the measurements reach steady state, which in turn enables constructing a system of nonlinear equations that is solved to come up with the desired initialization. The authors also compute the CRLB by using results from sensor array processing [91]. Using the previous framework, Jeremic and Nehorai in [92] tackle the problem of detecting and localizing landmines using a chemical sensor array. The authors also optimize the number and positions of sensors based on the CRLB. The same authors apply the previous ideas but in the context of ocean floor monitoring [93]. However the GLRT detection is used as the optimization criteria under the assumption that the medium parameters and sources locations are known. Zhao and Nehorai [94]

3.1. INTRODUCTION

investigated the detection and localization of a moving biochemical source. The open environment was modeled as a semi-infinite medium with unknown diffusivity. Maximum likelihood localization of a diffusive source is presented in [95] where the sensors use binary observations to preserve bandwidth. Matthes et. al. [96] on the other hand, proposed localizing a diffusive source via a two-step algorithm. In the first step the distances between the sensor and the source are estimated and in the second step the intersection of the previous distances is computed to find the final localization estimate. Localization in realistic environments was considered in [97] where a stochastic diffusion framework is used. A sequential version of the latter work was developed in [98].

Another approach for localizing a diffusive source is by using moving sensors instead of stationary ones. Porat and Nehorai [99], based on their previous work [90], proposed using a single moving sensor that follows a path set according to the CRLB gradient descent instead of the conventional concentration gradient. This work was extended to the case of multiple moving sensors in [100]. The previous algorithm is applied in [101], in which a group of *robotic fish* perform aquatic profiling.

The energy efficient distributed estimation method was introduced in Zhao and Nehorai [102] based on a distributed incremental implementation of the Gauss-Newton algorithm along with sensor selection based on maximizing the CRLB. The suggested algorithm though, was shown to resemble the extended Kalman filter, hence, the sensors exchange local statistics that involve a covariance-like matrix. A recent improvement on the previous work was given in [103] in which the initialization procedure was actually provided, in contrast to the latter work. The authors suggested solving the system of nonlinear equations of the steady-state measurement by semi-definite relaxation. Zhao and Nehorai also used distributed estimation but in a Bayesian framework in [104]. The belief (posterior pdf) is sequentially exchanged between the sensors. However, to preserve bandwidth only a few parameters that describe the belief are sent.

However, the work in papers [90]-[101] does not comply with constraints C1 and C2, since it assumes that the SNs can send all the raw data to the FC. This consumes considerable amount of bandwidth and energy, which is not tolerable in WSNs operation. Furthermore, it does not adhere to constraint C3 as well because straight forward MLE implementation is computational expensive. As is known, the success of MLE-based localization/estimation hinges on the choice of the initialization. The way it is computed in the reviewed literature above does not comply with the time-critical constraint C4, since steady state is required. Although the work [102]-[104] is energy efficient, it also suffers from the previous issue because it

3.1. INTRODUCTION

employs the same initialization technique.

3.1.2.2 Dynamical Parametric Model

In this approach, the PDE is used to formulate a discrete state-space model. The dynamics are obviously nonlinear, and therefore a linearization is often employed to be used with linear filtering, such as Kalman filter (KF) and its extended version (EKF), or nonlinear filtering such as Bayesian or particle filtering.

In [105, 106] the authors investigated the estimation of the PDE coefficient such as the diffusivity, wind speed, and dispersion for the purposes of field construction. The PDE was temporally and spatially discretized by using difference operators in order to formulate a discrete state space model. Then the PDE's parameters are estimated using an EKF. Fox et. al. [107] addressed the problem of detection and localization of a diffusive source by linearized approximation of the state-space equations describing the evolution of the plume [108]. The authors use nonlinear least squares estimation (NLSE) to find the parameters of the system and then plug them into the likelihood function to perform detection. The state-space model used above also needs a good initial value to linearize the model around it. However, this issue is not discussed in any of the above papers, and so it is not clear if constraint C4 is satisfied.

Sawo et. al. [109] solved the same problem using modal analysis (basis expansion such as Galerkin's method) instead of finite difference to formulate the state space model. Then the EKF is used to estimate the PDE's parameters. The authors also used a hybrid density filter [110, 111] and sliced Gaussian Mixture Filter [112, 113] to solve the problem, which are shown to be better than the EKF. Sundhar and Veeravalli [114] discussed localizing a diffusive source and tracking its intensity. The Karhunen-Loeve Galerkin method was used to convert the PDE to a discrete state space model that is solved via recursive prediction error technique, where the processing is performed centrally. An incremental distributed version was provided in the paper [115]. However, using Galerkin's method has its drawbacks, mainly the fact that in order to get high accuracy, a considerable number of terms must be included in the series expansion. This would increase the complexity of the system and hence violate constraint C3.

3.1. INTRODUCTION

3.1.2.3 Basis Representation Model

In this type of modeling, the diffusion field is represented by a linear combination of basis functions that are parametrized by the source parameters. The authors in [116] investigated the localization of diffusive point sources and the estimation of their release rates via representation by complex exponential eigenfunctions. The PDE now can be formulated into a system of linear equation that is used along with the annihilation filter [117] to find the location and the release rates of the sources. The same idea was used to find the release rates of smoke stakes in [118] using a number of chemical sensors for the purpose of environmental monitoring.

Although attractive from a computational point of view, this approach depends on the assumptions that the boundary condition is periodic, which might not apply to realistic scenarios.

3.1.2.4 Convolutional Model

Here the diffusion measurements are modeled as the convolution of a discretized kernel with the source signal. The localization/estimation problem is known as an inverse problem, which is well studied [119]. The inverse problem often boils down to the solution of a nonlinear estimation problem. However, such problems in this context are notoriously ill-posed. Hence the use of regularization techniques is needed. This framework was used to find the release rate in [120] and the location [121] of a polluting source under the atmospheric model in [108]. The same framework is used to identify and track chemical, biological, radiological, nuclear and explosive (CBRNE) sources in [122].

The issue of initialization in the NLSE is not addressed here. Furthermore, finding the optimal regularization parameter cannot be done in real-time, thus violating constraint C4.

3.1.3 Chapter Proposal and Contributions

In the light of the reviewed literature and the operational WSN constraints mentioned earlier, we investigate the problem of distributed estimation of a diffusive target in the context of the DDLE algorithm, which is part (II) highlighted in Fig. 3.1.1. The following contributions in the above problem are presented in this chapter:

1. We find an analytical expression for the CRLB of the desired parameters es-

3.1. INTRODUCTION

timates in a randomly deployed WSN using tools from stochastic geometry. These results are fairly general and can be applied to other physical phenomena.

2. As a part of the DDLE, the segmented distributed localization and estimation (SDLE) is introduced as an energy, bandwidth, and computation efficient framework in lieu of conventional algorithms. The main idea behind the SDLE is to introduce intermediate parameters into the primary model that enables the decoupling of the estimation problem into a local estimation problem, solved at the SN, and a global problem solved at the FC. This reduces the complexity of the system via solving a group of lower dimensional problems. Moreover, the estimated intermediate parameters are sent to the FC instead of the raw data to save energy and bandwidth. The SLDE builds on the distributed detection discussed in Chapter 2 and the assumption that the data sent to the FC is received successfully via the mechanism in Chapter 4. Three SDLE methods are presented:

- a) **NLSE:** NLSE is used to solve the local and global problems. The local problem is proved to be convex, and so no initialization is needed. The global localization problem can be easily initialized via the centroid method. In both previous problems, gradient-based methods can be used to find a solution. Alternatively, the grid search can be used to find the estimates in a way that is shown to be simpler than the gradient-based methods.
- b) **Hyper-Spherical Intersection :** Local problems are solved via NLSE, but the global problem is formulated as a hyper-spherical intersection (HSI) problem [123]. The HSI is ultimately solved by linear least squares (LLS). Although the LLS has a closed form solution, it is known to be sensitive to outliers. Hence, a robust HSI is proposed as well.
- c) **Hybrid-SDLE:** In this algorithm, GS is used for the local problem and NLSE is used for the global problem.

3.1.4 Chapter Outline

This chapter is organized as follows: Section 3.2 describes the physical and statistical models considered here. Then, the optimal maximum likelihood distributed

3.1. INTRODUCTION

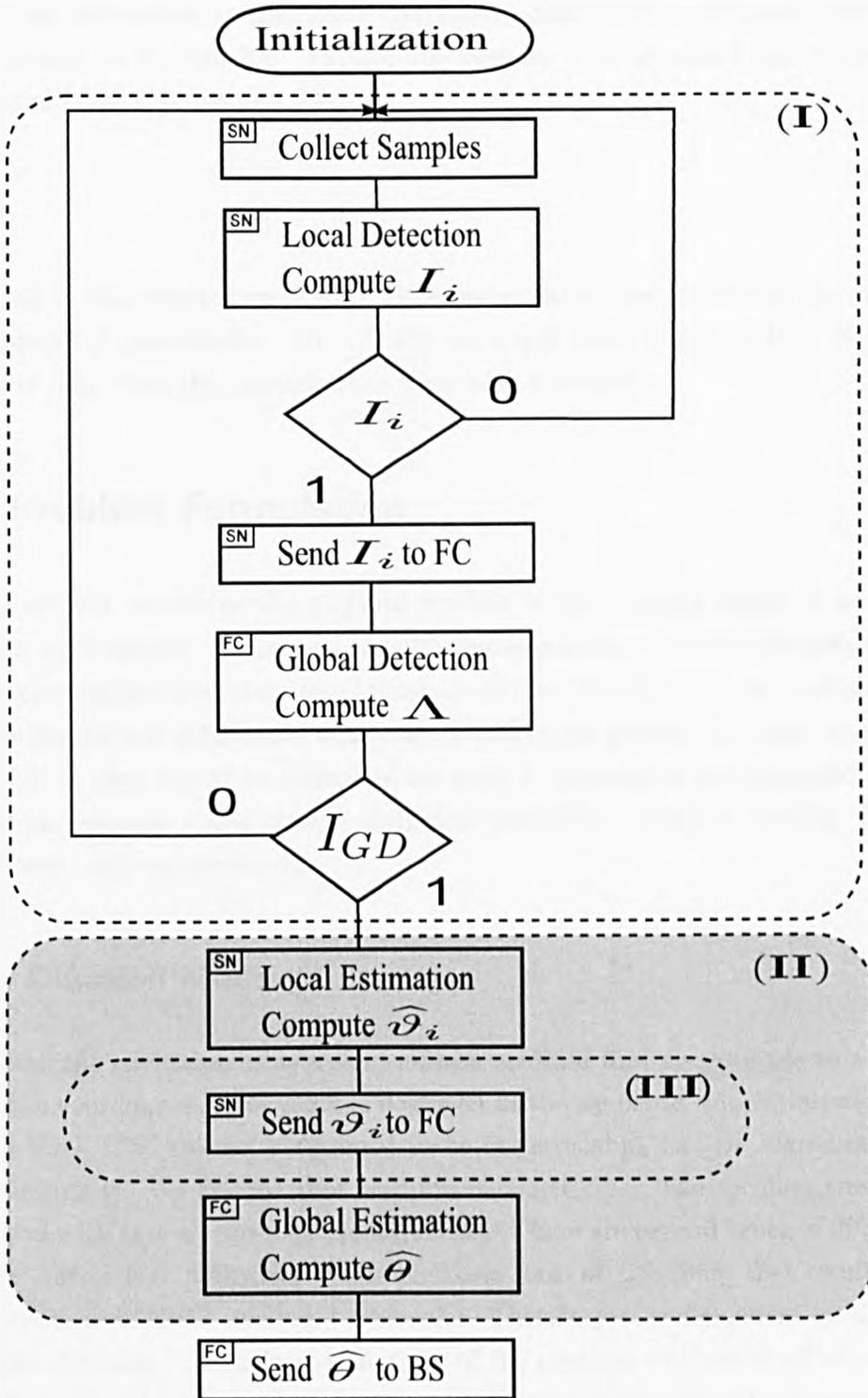


Figure 3.1.1: Distributed Detection, Localization, and Estimation (DDLE). In this chapter, section (II) is investigated.

3.2. PROBLEM FORMULATION

estimation is discussed in Section 3.3. The Cramer Rao lower bounds of the estimates are derived in Section 3.4. In Section 3.5 the segmented DLE is proposed as an efficient alternative to maximum likelihood. Simulation results and discussion are presented in Section 3.6. Finally, the chapter is summarized and concluding remarks are stated in Section 2.8.

Note

Note that in this chapter we slightly alter the notation used earlier for the purpose of simplicity of presentation. We will not use upper cases to denote RVs. However, it will be clear from the context when a variable is random.

3.2 Problem Formulation

In this section we derive the physical models of the target's signal in an open, outdoor, environment. We present the diffusion signal model in an open environment with corresponding boundary and initial conditions that describe the system. This leads to the partial differential equations (PDEs) that govern the signal transport. The PDE is then solved to formulate an analytic expression parametrized by the desired parameters. Using that, a statistical parametric model is derived. Finally, the system model is presented.

3.2.1 Diffusion Model

We model the environment as a semi-infinite medium that corresponds to $z > 0$ in Cartesian coordinates. The WSN is deployed in the xy -plane, which represents the ground level. The ground is assumed to be impermeable, i.e., the signal does not pass through it. We assume that medium is source free. The medium transports the signal with rate κ known as the *diffusivity*. There are several types of diffusivity however. *Molecular diffusivity* [55] is the basic form of diffusivity that results from the random walk of the medium's molecules. Therefore, it is dependent on the temperature, pressure, and molecular features of the medium and has small values. On the other hand, the *eddy diffusivity* is the transport resulting from having turbulent air currents, which is dominant in outdoor environments. The eddy diffusivity is dependent on the environmental conditions, hence it would be reasonable to assume it to be unknown [90]. This diffusivity is consequently space and time variant, i.e., it has different spatial components $\kappa(\mathbf{x}, t)$, where $\mathbf{x} = (x, y, z)^T$ is a point in $z > 0$

3.2. PROBLEM FORMULATION

(with some abuse of notation). However, as the eddy diffusivity is very complicated to model in theory, we adopt the methodology used in [90] and [94] where the molecular diffusion model is used with diffusivity values analogous to the eddy diffusivity.

The intruding target is a point source at location $\mathbf{x}_t = (x_t, y_t, z_t)^T$ that initiates a transport of matter at an arbitrary time, say τ . This matter can be a biological material, chemical material, or heat for example. The target's release rate is modeled as

$$\mu(\mathbf{x}, t) = (\mu + \tilde{\mu}(t)) \delta(\|\mathbf{x} - \mathbf{x}_t\|) \quad (3.2.1)$$

where μ is the constant or mean component, $\tilde{\mu}(t)$ is the time-varying component, and $\delta(\cdot)$ is the Kronecher delta function.

The target's matter transport signal obeys the following diffusion law

$$\frac{\partial}{\partial t} s(\mathbf{x}, t) = -\nabla^2 (\kappa(\mathbf{x}, t) s(\mathbf{x}, t)) \quad (3.2.2)$$

where $\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the spatial second derivative. The impermeable ground is represented by the following boundary condition

$$\frac{\partial}{\partial z} s(\mathbf{x}, t) = 0, \forall x, y : \mathbf{x} = (x, y, 0)^T. \quad (3.2.3)$$

The following

$$s(\mathbf{x}, \tau^+) = \mu(\tau^+) \delta(\|\mathbf{x} - \mathbf{x}_t\|) \quad (3.2.4)$$

$$s(\mathbf{x}, \tau^-) = 0 \quad (3.2.5)$$

represent the initial conditions, which includes the target's signal.

The PDE (3.2.2) is difficult to solve in its current form. Hence, we impose some reasonable assumptions that simplify the solution. First we assume that $\kappa(\mathbf{x}, t)$ does not significantly change in the period of time observed by the SNs. Furthermore, the detecting SNs are spatially close to each other, and so κ does not experience a large change spatially. Hence, we can assume that the diffusivity is constant across the detecting SNs within the sampling period, i.e., $\kappa(\mathbf{x}, t) \approx \kappa$. In this case (3.2.2) reduces to

3.2. PROBLEM FORMULATION

$$\frac{\partial}{\partial t} s(\mathbf{x}, t) = -\kappa \nabla^2 s(\mathbf{x}, t). \quad (3.2.6)$$

Now we can solve for the above boundary and initial conditions by first finding the corresponding Green function [124], say $G(\mathbf{x}, t)$, and then computing the final solution as

$$s(\mathbf{x}, t) = \int_{-\infty}^{\infty} \int_{\mathbb{R}^3} G(\mathbf{x}, t) \mu(\mathbf{x}, t) d\mathbf{x} dt. \quad (3.2.7)$$

However, since we are dealing with a semi-infinite medium here, it would be easier to find the solution for the infinite medium case and use the method of mirrors to arrive at the desired solution [94]. Thus, the Green function for the infinite case is found by solving (3.2.6) without the boundary condition (3.2.3) for an impulse source in space and time that gives [124]

$$G_{inf}(\mathbf{x}, t) = \frac{1}{8 [\pi \kappa (t - \tau)]^{\frac{3}{2}}} \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_t\|^2}{4\kappa(t - \tau)}\right). \quad (3.2.8)$$

Plugging (3.2.8) into (3.2.7) yields the final solution. In order to simplify matters, we restrict our analysis to the mean component of (3.2.1), and so we arrive at [90]

$$\begin{aligned} s_{inf}(\mathbf{x}, t) &= \int_{\tau}^t \frac{\mu}{8 [\pi \kappa (t' - \tau)]^{\frac{3}{2}}} \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}_t\|^2}{4\kappa(t' - \tau)}\right) dt' \\ &= \frac{\mu}{4\pi\kappa \|\mathbf{x} - \mathbf{x}_t\|} \operatorname{erfc}\left(\frac{\|\mathbf{x} - \mathbf{x}_t\|}{2\sqrt{\kappa(t - \tau)}}\right). \end{aligned} \quad (3.2.9)$$

The solution for the semi-infinite case now is

$$s(\mathbf{x}, t) = \frac{\mu}{4\pi\kappa \|\mathbf{x} - \mathbf{x}_t\|} \operatorname{erfc}\left(\frac{\|\mathbf{x} - \mathbf{x}_t\|}{2\sqrt{\kappa(t - \tau)}}\right) + \frac{\mu}{4\pi\kappa \|\mathbf{x} - \mathbf{x}_t\|} \operatorname{erfc}\left(\frac{\|\mathbf{x} - \mathbf{x}'_t\|}{2\sqrt{\kappa(t - \tau)}}\right) \quad (3.2.10)$$

where $\mathbf{x}'_t = (x_t, y_t, -z_t)$.

If the SNs are assumed to be deployed on the flat xy -plane, then the two terms in the above expression are identical, since the distances are symmetrical. Consequently the above can be simplified to

3.2. PROBLEM FORMULATION

$$s(\mathbf{x}, t) = \frac{\mu}{2\pi\kappa \|\mathbf{x} - \mathbf{x}_t\|} \operatorname{erfc} \left(\frac{\|\mathbf{x} - \mathbf{x}_t\|}{2\sqrt{\kappa(t - \tau)}} \right). \quad (3.2.11)$$

The above signal model will be used for the rest of this chapter.

3.2.2 System Model

Target

We assume that an intruder target is a diffusive point source located at an unknown and random location \mathbf{x}_t in \mathbb{R}^3 . The point source assumption is valid when the distance between the target and the SNs are much larger than the target's physical dimensions. The target leaves a signature that is governed by the above model (3.2.11). It is characterized by the unknown parameter vector $(\mu, \mathbf{x}_t^T, \tau)$, where, as stated earlier, μ is the release rate and τ is the entry time.

Sensor node

SNs are deployed in a sensing field \mathcal{F} where $\mathcal{F} \subset \mathbb{R}^2$. They are located at points $\{\mathbf{x}_i\}$ that are i.i.d according to a uniform distribution. The SNs are equipped with a sensing device able to measure diffusion related signals. Every SN collects N samples at a sampling frequency of $1/T_s$. The measurements are corrupted by AWGN with known variance σ_s^2 . The noise is assumed to be spatially and temporally independent, i.e.,

$$\mathbb{E}[\mathbf{w}_i^T \mathbf{w}_j] = \begin{cases} N\sigma_s^2, & i = j \\ 0, & i \neq j \end{cases} \quad (3.2.12)$$

where $\mathbf{w}_i = (w_i[0], \dots, w_i[N-1])^T$ are the noise samples collected by the SN located at \mathbf{x}_i . The SN also has an energy detector (to perform local detection) and a local processing unit to extract local information from the measured signal.

Communication channel

For the purposes in this chapter, we assume that there exists a perfect communication channel between every SN and the FC. In other words, the FC correctly receives any messages from the SNs. This goal can be achieved by proper channel coding

3.3. CLASSICAL DISTRIBUTED ESTIMATION

and the use of a MAC protocol. The latter is discussed in Chapter 4.

WSN

A network of SNs is deployed to monitor and collect measurements of any occurring diffusion event in the sensing field. The WSN is modeled by a PPP $\Phi = \{\mathbf{x}_i\}$, $i \in \mathbb{N}$ with intensity λ where $\mathbf{x}_i \in \mathbb{R}^2$ is the location of the i th SN. This implies that the SNs are deployed in the xy -plane. The target location \mathbf{x}_t is obviously not collocated with any SNs, i.e., $\mathbf{x}_t \notin \Phi$.

3.2.3 Distributed Estimation Problem

In this chapter, the problem of distributed estimation of the diffusion model presented above is considered in the DDLE framework. More particularly, the i th SN collected the measurements

$$f_i[n] = g[n; \mathbf{x}_i, \boldsymbol{\theta}] + w_i[n] \quad (3.2.13)$$

where $n = 0, \dots, N - 1$, $\boldsymbol{\theta} = (\mu, \mathbf{x}_t^T, n_\tau, \kappa)^T$ is the vector of target and medium parameters, and

$$g[n; \mathbf{x}_i, \boldsymbol{\theta}] = \frac{\mu}{2\pi\kappa \|\mathbf{x}_i - \mathbf{x}_t\|} \operatorname{erfc} \left(\frac{\|\mathbf{x}_i - \mathbf{x}_t\|}{2\sqrt{\kappa(n - n_\tau)T_s}} \right). \quad (3.2.14)$$

We wish to find a resource efficient distributed estimation algorithm to estimate $\boldsymbol{\theta}$ with acceptable accuracy. To this end, the WSN should have performed distributed detection and arrived at a positive decision about the target's presence, then proceeded to the distributed estimation stage as suggested by the DDLE algorithm (see Fig. 3.1.1).

3.3 Classical Distributed Estimation

In this section, the classical centralized estimation problem is discussed in order to set the stage for the distributed estimation problem in a random network. First, we conveniently reformulate the measurement model (3.2.13) for convenience to

$$\mathbf{f}_i = \mu \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) + \mathbf{w}_i \quad (3.3.1)$$

3.3. CLASSICAL DISTRIBUTED ESTIMATION

where $\boldsymbol{\psi} = (\mathbf{x}_i^T, \tau, \kappa)$, $\mathbf{f}_i = (f_i[0], \dots, f_i[N-1])$, and $\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi})$ is the vector format of (3.2.14) divided by μ . Note however, that (3.3.1) can also represent different physical phenomena such as wave propagation.

To find the optimal estimates of the target parameters the joint likelihood function of the SNs measurements must be first formulated. Let the number of SNs in the network be M with known locations (the \mathbf{x}_i 's are known). Then the joint likelihood function of their measurements is

$$\begin{aligned} p(\mathbf{f}_0, \dots, \mathbf{f}_{M-1}; \boldsymbol{\theta}) &= \prod_{i=0}^{M-1} \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left(-\frac{\|\mathbf{f}_i - \mu\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi})\|^2}{2\sigma_s^2}\right) \\ &= (2\pi\sigma_s^2)^{-\frac{M}{2}} \exp\left(-\sum_{i=0}^{M-1} \|\mathbf{f}_i - \mu\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi})\|^2\right). \end{aligned} \quad (3.3.2)$$

Define the log-likelihood function (LLF) as

$$l(\boldsymbol{\theta}) = -\frac{1}{2} \sum_{i=0}^{M-1} \|\mathbf{f}_i - \mu\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi})\|^2 \quad (3.3.3)$$

where we have ignored the term independent of $\boldsymbol{\theta}$. The optimal maximum likelihood estimate of $\boldsymbol{\theta}$ is the solution of

$$\begin{aligned} \hat{\boldsymbol{\theta}}_{\text{MLE}} &= \arg \max_{\boldsymbol{\theta}} l(\boldsymbol{\theta}) \\ &\equiv \arg \min_{\boldsymbol{\theta}} \frac{1}{2} \sum_{i=0}^{M-1} \|\mathbf{f}_i - \mu\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi})\|^2. \end{aligned} \quad (3.3.4)$$

Let us group the measurement vectors into the matrix $\mathbf{F} = [\mathbf{f}_0 \cdots \mathbf{f}_{M-1}]$, and the transfer vectors into the transfer matrix $\mathbf{H}(\boldsymbol{\psi}) = [\mathbf{h}(\mathbf{x}_0, \boldsymbol{\psi}) \cdots \mathbf{h}(\mathbf{x}_{M-1}, \boldsymbol{\psi})]$. Then (3.3.3) can be formulated as

$$\begin{aligned} l(\boldsymbol{\theta}) &= \frac{1}{2} \text{tr} \{ (\mathbf{F} - \mu\mathbf{H}(\boldsymbol{\psi})) (\mathbf{F} - \mu\mathbf{H}(\boldsymbol{\psi}))^T \} \\ &= \frac{1}{2} \|\mathbf{F} - \mu\mathbf{H}(\boldsymbol{\psi})\|_F^2 \end{aligned} \quad (3.3.5)$$

where $\text{tr}\{\cdot\}$ is the matrix trace and $\|\cdot\|_F$ is the Frobenius norm. It can be readily shown that the maximum likelihood estimates are

3.4. CRAMER RAO LOWER BOUND

$$\hat{\boldsymbol{\psi}}_{\text{MLE}} = \arg \max_{\boldsymbol{\psi}} \frac{\|\mathbf{F}\mathbf{H}^T(\boldsymbol{\psi})\|_F^2}{\|\mathbf{H}\mathbf{H}^T(\boldsymbol{\psi})\|_F} \quad (3.3.6)$$

and

$$\hat{\mu} = \frac{\|\mathbf{F}\mathbf{H}^T(\hat{\boldsymbol{\psi}})\|_F^2}{\|\mathbf{H}\mathbf{H}^T(\hat{\boldsymbol{\psi}})\|_F^2}. \quad (3.3.7)$$

The MLE is known to be asymptotically consistent and efficient given that some regularity conditions are satisfied [125], i.e., the MLE is asymptotically unbiased and achieves the CRLB. However, when a random network is considered, the LLF in (3.3.3) becomes more complex since both the number of SNs and their locations are random. Hence, the MLE in this case is also complicated. Thus, we resort to finding the CRLB of the MLE in the random network case.

3.4 Cramer Rao Lower Bound

In this section, the CRLB is derived for a WSN that is modeled by a PPP. Moreover, the use of censoring is also included in the discussion.

The estimation/localization performance is characterized by computing the CRLB bound and comparing it to the mean square errors of the actual estimates. The CRLB is defined as

$$\text{CRLB}(\boldsymbol{\theta}) = [\mathcal{I}(\boldsymbol{\theta})]^{-1} \quad (3.4.1)$$

where

$$\mathcal{I}(\boldsymbol{\theta}) = -\mathbb{E} \left[\frac{\partial^2 \mathcal{l}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T \boldsymbol{\theta}} \right] \quad (3.4.2)$$

is the Fisher information matrix (FIM).

Obviously the CRLB is inversely proportional to the FIM. Thus, in order to get an insight into the CRLB, the FIM should be calculated. From (3.4.2), the FIM is the expectation of negative curvature of the LLF at $\boldsymbol{\theta}$. In other words, it is a measure of the sharpness of the LLF at the true parameters value. A LLF with high curvature means that it is sharply centered around the true parameters value, and so there is a low uncertainty about the $\boldsymbol{\theta}$ value. By contrast, a LLF with low

3.4. CRAMER RAO LOWER BOUND

curvature means that the LLF is spread out, which in turn implies that there is high uncertainty about θ . Note however, that the FIM cannot be negative since the curvature at θ , which is the global minimum of the LLF, is always negative. Having said that, the CRLB now represents the uncertainty about the parameters, which is manifested as the estimation variance.

The FIM for a PPP modeled WSN is computed using (3.4.2) and (3.3.3) as follows

$$\mathcal{I}(\theta) = \mathbb{E} \left[\frac{1}{2\sigma_s^2} \frac{\partial^2}{\partial \theta^T \theta} \sum_{\mathbf{x}_i \in \Phi} \|\mathbf{f}_i - \mu \mathbf{h}(\mathbf{x}_i, \psi)\|^2 \right]. \quad (3.4.3)$$

The above can be simplified by the additive property of the FIM and the fact that the measurement noise and the network realization, Φ , is independent, and so

$$\mathcal{I}(\theta) = \mathbb{E}_{\Phi} \left[\sum_{\mathbf{x}_i \in \Phi} \frac{1}{2\sigma_s^2} \mathbb{E}_w \left[\frac{\partial^2}{\partial \theta^T \theta} \|\mathbf{f}_i - \mu \mathbf{h}(\mathbf{x}_i, \psi)\|^2 \right] \right] \quad (3.4.4)$$

$$= \mathbb{E}_{\Phi} \left[\sum_{\mathbf{x}_i \in \Phi} \mathcal{I}(\mathbf{x}_i, \theta) \right] \quad (3.4.5)$$

i.e., the ultimate FIM is the aggregation of the individual SN's information. In the case of censoring the SN, the FIM simply becomes

$$\mathcal{I}_{\text{cen}}(\theta) = \mathbb{E}_{\Phi_d} \left[\sum_{\mathbf{x}_i \in \Phi_d} \mathcal{I}(\mathbf{x}_i, \theta) \right]. \quad (3.4.6)$$

The FIM at a given SN is given by [94]

$$\mathcal{I}(\mathbf{x}_i, \theta) = \begin{pmatrix} \mathbf{h}^T(\mathbf{x}_i, \psi) \mathbf{h}(\mathbf{x}_i, \psi) & \mathbf{h}^T(\mathbf{x}_i, \psi) \mathbf{D}(\mathbf{x}_i, \psi) \\ \mathbf{D}^T(\mathbf{x}_i, \psi) \mathbf{h}(\mathbf{x}_i, \psi) & \mathbf{D}^T(\mathbf{x}_i, \psi) \mathbf{D}(\mathbf{x}_i, \psi) \end{pmatrix} \quad (3.4.7)$$

where

$$\mathbf{D}(\mathbf{x}, \psi) = \frac{\partial}{\partial \psi} \mathbf{h}(\mathbf{x}, \psi) \quad (3.4.8)$$

is the derivative matrix whose columns are $\frac{\partial}{\partial \psi_i} \mathbf{h}(\mathbf{x}, \psi_i)$. The full derivation though, is given in B.2.

3.4. CRAMER RAO LOWER BOUND

However, the computation of (3.4.5) or (3.4.6) is not straightforward since the FIM depends on the stochastic geometry of the network, nature of censoring if used, and the nonlinear dependence on θ .

Nonetheless, stochastic geometry tools can be used here to find an expression for the CRLB as shown in the following theorem.

Theorem 3.1. *Assume that the conventional regularity conditions for the CRLB are satisfied [125]. And let $P_d(\mathbf{x}, \gamma)$ be the censoring probability with threshold γ . Let the following conditions hold*

$$\mathbf{D}^T(\mathbf{x}, \boldsymbol{\psi})\mathbf{h}(\mathbf{x}, \boldsymbol{\psi}) \geq 0 \quad (3.4.9)$$

for all \mathbf{x} and $\boldsymbol{\psi}$, in addition to the existence of the following integrals

$$\tilde{\mathbf{h}}(\boldsymbol{\psi}, \gamma) = \int_{\mathbb{R}^2} \|\mathbf{h}(\mathbf{x}, \boldsymbol{\psi})\|^2 P_d(\mathbf{x}, \gamma) d\mathbf{x} \quad (3.4.10)$$

$$\tilde{\mathbf{G}}(\boldsymbol{\psi}, \gamma) = \int_{\mathbb{R}^2} \mathbf{D}^T(\mathbf{x}, \boldsymbol{\psi})\mathbf{h}(\mathbf{x}, \boldsymbol{\psi}) P_d(\mathbf{x}, \gamma) d\mathbf{x} \quad (3.4.11)$$

$$\tilde{\mathbf{D}}(\boldsymbol{\psi}, \gamma) = \int_{\mathbb{R}^2} \|\mathbf{D}(\mathbf{x}, \boldsymbol{\psi})\|^2 P_d(\mathbf{x}, \gamma) d\mathbf{x} \quad (3.4.12)$$

for all \mathbf{x} and $\boldsymbol{\psi}$ values in addition to $\tilde{\mathbf{h}}(\boldsymbol{\psi}, \gamma) \neq 0$. Then the CRLB for the $\boldsymbol{\theta} = (\mu, \boldsymbol{\psi}^T)^T$ is

$$\text{CRLB}(\boldsymbol{\psi}) = \frac{\sigma_s^2}{\lambda\mu^2} \left(\tilde{\mathbf{D}}(\boldsymbol{\psi}, \gamma) - \frac{\tilde{\mathbf{G}}^T(\boldsymbol{\psi}, \gamma)\tilde{\mathbf{G}}(\boldsymbol{\psi}, \gamma)}{\tilde{\mathbf{h}}(\boldsymbol{\psi}, \gamma)} \right)^{-1} \quad (3.4.13)$$

$$\text{CRLB}(\mu) = \frac{\sigma_s^2}{\lambda} \left(\tilde{\mathbf{h}}(\boldsymbol{\psi}, \gamma)\mathbf{I} - \tilde{\mathbf{G}}(\boldsymbol{\psi}, \gamma)\tilde{\mathbf{D}}^{-1}(\boldsymbol{\psi}, \gamma)\tilde{\mathbf{G}}^T(\boldsymbol{\psi}, \gamma) \right)^{-1}. \quad (3.4.14)$$

Proof. See Appendix B.3. □

The bounds (3.4.13) and (3.4.14) coincide with the results given in [94] if the number of SNs is known and they are not censored, i.e., $P_d(x) = 1$. More importantly, the previous bound shows inverse proportionality to the deployment density of the SNs, λ . In other words, the CRLB improves when λ is increased. This is attributed to having more SNs providing information, which improves the estimation.

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

However, the influence of γ on the CRLB is not explicitly clear in the theorem. To understand it, we need to look first at the non-censored case. Intuitively, SNs farther away from the target are expected to provide less information when compared to closer SNs. In other words

$$\|\mathcal{I}(\mathbf{x}_i, \boldsymbol{\theta})\| \propto \frac{1}{\|\mathbf{x}_i - \mathbf{x}_t\|} \quad (3.4.15)$$

where the norm in the left hand side is a general matrix norm. Hence, it is expected that SNs located outside a disk with some radius R would provide negligible information. If censoring is used though, only a subset of the SNs will provide information for the estimation to the FC. A SN is a member of this subset with probability $P_d(\mathbf{x}, \gamma)$, which is the local detection probability. Apparently, $P_d(\mathbf{x}_i)$ (we drop γ for convenience) is also inversely proportional to $\|\mathbf{x}_i - \mathbf{x}_t\|$. To simplify the discussion, let us approximate $P_d(\mathbf{x}_i)$ with the step function (see discussion in Section 2.3)

$$P_d(\mathbf{x}_i) = \begin{cases} 1, & \|\mathbf{x}_i - \mathbf{x}_t\| < r \\ 0, & \|\mathbf{x}_i - \mathbf{x}_t\| > r \end{cases} \quad (3.4.16)$$

where r is some distance that is inversely proportional to γ . In other words, the SN only detects the target, and consequently sends off information to the FC, with probability one if it is located inside a disk with radius r . Now if the local detection threshold is increased such that $r < R$, not all the “useful” information is sent to the FC. On the other hand, if $r > R$ then SNs with negligible information also send their data to the FC. In terms of the CRLB however, it is expected that decreasing γ will improve the CRLB to a certain point after which the CRLB will not change significantly.

3.5 Segmented Distributed Localization and Estimation

Although optimal, the MLE discussed earlier suffers from high communication and computational requirements that are undesirable for a realistic WSN. Therefore, we propose the SDLE algorithm that trades off performance with needed resources and complies with the time-critical constraints required by the application.

In this section we discuss the elements of SDLE and the implementation of its local and global estimation components. In both components, a nonlinear least squares problem is solved via iterative gradient based algorithms that are used for

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

estimation in addition to grid search methods. The global estimation is also solved via reformulation of the problem as a linear least squares problem. The grid search however, is shown to be more computationally efficient than the other methods.

3.5.1 Elements of SDLE

The philosophy of the SDLE is to decompose the MLE problem into decoupled local problems solved at the SNs and a global problem solved at the FC. The local estimation of intermediate parameters that are introduced to decouple the original problem (3.3.4) into simpler (lower dimension) local estimation problems. If the distributed detection stage results in a positive global detection (see Fig. 2.1.1), the local intermediate estimates are then sent to the FC for the second phase instead of sending the collected measurements, which considerably saves communication power and bandwidth. Chapter 4 discusses the communication side of the SDLE. In global estimation, the FC uses the received intermediate parameter estimates to estimate the desired target (and medium) parameters, which we call here the global estimates. This concept is illustrated in Fig. 3.5.1. Although this approach is resource efficient, it is not expected, however, to perform as well as the optimal approach described in Subsection 3.3 since the received intermediate parameters estimates used for global estimation might contain some errors. Thus, the SDLE trades-off estimation performance with network resource efficiency.

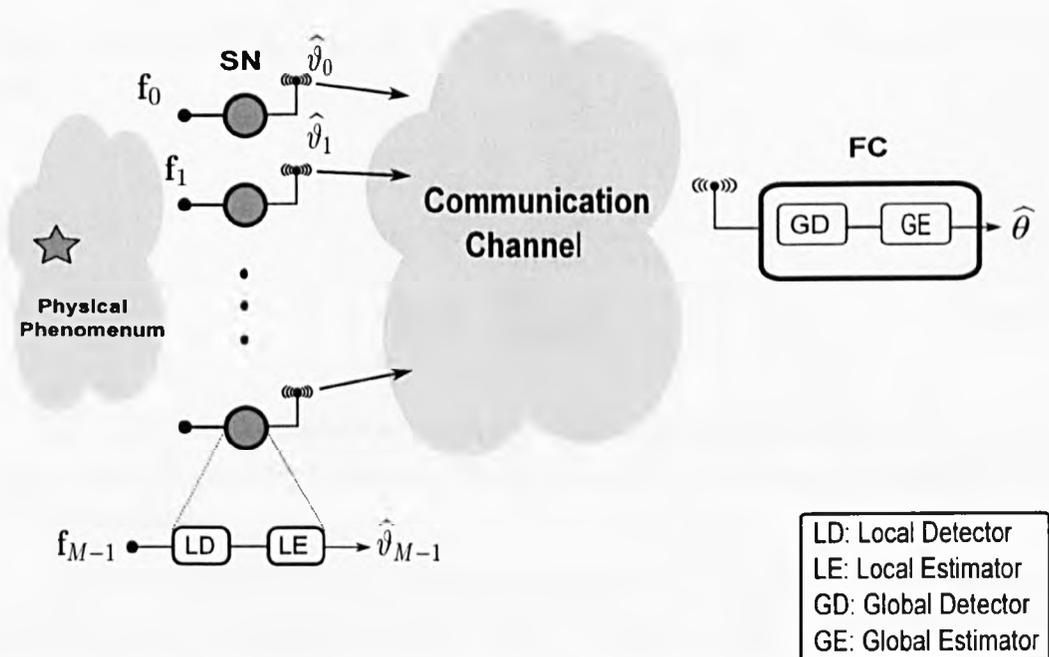


Figure 3.5.1: Segmented distributed localization and estimation system diagram. $\hat{\vartheta}_i$ is an intermediate local estimate and $\hat{\theta}$ is the global estimate of the desired parameters.

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

The SDLE relies on two main concepts. First, decomposition into local and global processing and the second is the reparameterization of the statistical model to facilitate the desired decomposition.

3.5.1.1 Reparameterization

At first glance, the estimation problem (3.3.4) can be decomposed into local estimation problems solved at the SNs. However, this cannot be done directly since estimating \mathbf{x}_t requires information from at least four different SNs [90]. In order to decouple (3.3.4) into locally independent problems, we introduce intermediate parameters to be estimated in lieu of the original parameters $\boldsymbol{\theta}$ by re-parameterizing the diffusion case

$$g[n; \mathbf{x}_i, \boldsymbol{\theta}] = \frac{\mu}{2\pi\kappa \|\mathbf{x}_i - \mathbf{x}_t\|} \operatorname{erfc} \left(\frac{\|\mathbf{x}_i - \mathbf{x}_t\|}{2\sqrt{\kappa(n - n_\tau)T_s}} \right) \quad (3.5.1)$$

into

$$\begin{aligned} g[n; \mathbf{x}_i, \boldsymbol{\vartheta}_i] &= a_i \operatorname{erfc} \left(\frac{b_i}{\sqrt{n - n_\tau}} \right) \\ &= a_i h[n, \boldsymbol{\varphi}_i] \end{aligned} \quad (3.5.2)$$

where the intermediate parameters now are $\boldsymbol{\vartheta}_i^T = (a_i, \boldsymbol{\varphi}_i^T)$, in which $\boldsymbol{\varphi}_i = (b_i, n_\tau)^T$ and

$$a_i = \frac{\mu}{2\pi\kappa \|\mathbf{x}_i - \mathbf{x}_t\|} \quad (3.5.3)$$

$$b_i = \frac{\|\mathbf{x}_i - \mathbf{x}_t\|}{2\sqrt{\kappa T_s}}. \quad (3.5.4)$$

Note that the parameters a_i and b_i have different values for every SN, in contrast to n_τ , which is constant. However, the estimate of n_τ is expected to vary from one SN to the other.

Another interesting interpretation of the above reparameterization is via using the *finite rate of innovation* concept [117]. In this context, the signal under consideration has limited innovation, hence it can be represented by a finite number of free parameters. Fortunately, this idea suits our case, and most naturally generated phenomena. The diffusion concentration signal in (3.5.1) can be completely

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

represented by the parameters a_i , b_i , and n_τ . Therefore, we can *code* the signal in (3.5.1) by the set estimates of the previous parameters, and send the code to save bandwidth.

3.5.1.2 Distributed Detection

The reparameterization above also enables the decoupling of the detection problem into local and global parts. In the local detection, the i th SN tests for

$$\mathcal{H}_0 : f_i[n] = w_i[n], n = 0, \dots, n_\tau \quad (3.5.5)$$

$$\mathcal{H}_1 : f_i[n] = a_i \operatorname{erfc} \left(\frac{b_i}{\sqrt{n - n_\tau}} \right) + w_i[n] \quad n = n_\tau + 1, \dots, N - 1. \quad (3.5.6)$$

However, the intermediate parameters are not known, hence the GLRT can be used. Although it is optimal, it leads to unnecessary processing if the target is not present. Therefore, we adopt the paradigm in Fig. 3.5.1, in which the ED is used because it requires no knowledge of ϑ_i . The ED computes the energy of the signal, i.e.,

$$\Xi(\mathbf{x}_i) = \sum_{n=0}^{N-1} f_i^2[n]. \quad (3.5.7)$$

The local decision is computed by comparing the above test statistic to the local threshold γ . In other words

$$I(\mathbf{x}_i) = \begin{cases} 1, & \Xi(\mathbf{x}_i) > \gamma \\ 0, & \text{Otherwise} \end{cases} \quad (3.5.8)$$

The false alarm and detection probability follows as

$$P_{fa} = Q_{\chi^2_N}(\gamma) \quad (3.5.9)$$

$$P_d(\mathbf{x}_i) = Q_{\frac{N}{2}} \left(\sqrt{\frac{\Xi(\mathbf{x}_i)}{\sigma_s^2}}, \sqrt{\gamma} \right) \quad (3.5.10)$$

respectively, where $Q_{\chi^2_N}(\cdot)$ is the tail probability of a central χ^2 distribution with N degrees of freedom and $Q_{\frac{N}{2}}(\cdot, \cdot)$ is the generalized Marcum Q-function. Recall that γ is set according to a specified P_{fa} found from (3.5.9).

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

Now the SNs with positive local decisions send their decisions to the FC where the global detection occurs according to the CR

$$\Lambda_{CR} = \sum_{\mathbf{x}_i \in \Phi} I(\mathbf{x}_i) \quad (3.5.11)$$

and the global decision takes the form

$$I_{GD} = \begin{cases} 1, & \Lambda_{CR} \geq \Upsilon \\ 0, & \text{Otherwise} \end{cases} \quad (3.5.12)$$

where Υ is the global detection threshold that is set according to a certain global false alarm probability, P_{FA} . The global false alarm and detection probabilities are given as

$$P_{FA} = \sum_{m=\Upsilon}^{\infty} \frac{(\bar{\Omega}_0)^m}{k!} e^{-\bar{\Omega}_0} \quad (3.5.13)$$

$$P_D = \sum_{m=\Upsilon}^{\infty} \frac{(\bar{\Omega}_1)^m}{k!} e^{-\bar{\Omega}_1} \quad (3.5.14)$$

where

$$\bar{\Omega}_0 = P_{fa} |\mathcal{F}| \quad (3.5.15)$$

$$\bar{\Omega}_1 = \int_{\mathcal{F}} P_d(\mathbf{x}) d\mathbf{x}. \quad (3.5.16)$$

Note however, that knowledge of ϑ_i is not needed to compute Υ but it is needed to compute P_D through (3.5.10) and (3.5.14). To address this issue, we suggest assigning a prior distribution to ϑ_i and use theorem 2.2 to find P_D . This value will lower bound the actual performance.

3.5.1.3 Local and Global Processing

Using the reparameterization in (3.5.2), (3.5.3), and (3.5.4) the original problem (3.3.4) can be decoupled into simpler (lower dimension) local estimation problems solved individually at each detecting SN. Thus, the MLE problem (3.3.4) can be reformulated in terms of the intermediate parameters as

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

$$\min_{\{\boldsymbol{\vartheta}_0, \dots, \boldsymbol{\vartheta}_{M-1}\}} \frac{1}{2} \sum_{i=0}^{M-1} \|\mathbf{f}_i - a_i \mathbf{h}(\mathbf{x}_i, \boldsymbol{\varphi}_i)\|^2 \quad (3.5.17)$$

where M is the number of active SNs, $\boldsymbol{\vartheta}_i = (a_i, \boldsymbol{\varphi}_i^T)^T$ and $\boldsymbol{\varphi}_i = (b_i, n_\tau)^T$ as mentioned earlier. Now the above estimation problem can be decomposed into the following M problems solved individually at the detecting SNs

$$\min_{(a_i, \boldsymbol{\varphi}_i)} \|\mathbf{f}_i - a_i \mathbf{h}(\mathbf{x}_i, \boldsymbol{\varphi}_i)\|^2. \quad (3.5.18)$$

The solution of the MLE in (3.5.18) above is given in Subsection 3.5.2. The estimated intermediate parameters $(\hat{a}_i, \hat{\boldsymbol{\varphi}}_i)$ are then transmitted to the FC in order to solve for the ultimate global parameters $\boldsymbol{\theta}$ as shown in Fig. 3.5.1. This global estimation can be regarded as a fusion process of the local estimates to produce the desired result. The manner in which the fusion is carried out is not particularly specified by (3.3.4) as it is the case for local estimation. Hence, the fusion problem can be approached with more freedom. In this work, we generally adopt the least squares method. Given $\hat{\boldsymbol{\vartheta}}_i$ and from the definition (3.5.4), the least squares approach can be used to find the position of the target and the medium's diffusivity, i.e.,

$$(\hat{\mathbf{x}}_t, \hat{\kappa}) = \arg \min_{\mathbf{x}_t, \kappa} \sum_{i=0}^{M-1} \left(\hat{b}_i - \frac{\|\mathbf{x}_i - \mathbf{x}_t\|}{2\sqrt{\kappa T_s}} \right)^2 \quad (3.5.19)$$

where T_s is the sampling period. The above is actually a nonlinear least squares problem and does not have a closed form solution, however it can be solved iteratively by a variety of methods, which are discussed later on in this chapter. Note that the previous estimator implies that the FC has knowledge of the transmitting SNs. So using linear least squares to find the entry time estimate ($\hat{\tau}$) leads to the estimator

$$\hat{\tau} = \frac{T_s}{M} \sum_{i=0}^{M-1} \hat{n}_\tau^i \quad (3.5.20)$$

where \hat{n}_τ^i is the entry time sample estimate of the i th SN. The above estimator is actually a moving average filter. Similarly, the release rate (μ) is estimated according to

$$\hat{\mu} = \frac{2\pi\hat{\kappa}}{M} \sum_{i=0}^{M-1} \hat{a}_i \|\mathbf{x}_i - \hat{\mathbf{x}}_t\|. \quad (3.5.21)$$

It is expected though that $\hat{\mu}$ would not be accurate since it depends on several other estimates. We will now present local and global processing techniques to solve

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

(3.5.18) and (3.5.19) respectively.

3.5.2 Local Estimation Techniques

In the following we discuss techniques to solve the local estimation problem (3.5.18). First, we investigate a gradient-based search approach and prove that it can be formulated as a convex problem, which ensures reaching an acceptable solution. However, it turns out that the latter technique might be computationally demanding in the context of WSN. Therefore, we suggest using a grid search technique that turns out to be a simple correlator with pre-stored values.

3.5.2.1 Gradient-Based Search

To solve (3.5.18), we employ the variable projection (VP) method [126]. The VP uses the fact that a subset of parameters exist linearly in the cost function to eliminate those parameters from the problem. The (a) parameter can be estimated using linear least squares as

$$\hat{a}_i = \frac{\mathbf{h}^T(\hat{\boldsymbol{\varphi}}_i)\mathbf{f}_i}{\mathbf{h}^T(\hat{\boldsymbol{\varphi}}_i)\mathbf{h}(\hat{\boldsymbol{\varphi}}_i)} \quad (3.5.22)$$

where we dropped the \mathbf{x}_i argument for notational convenience. $\hat{\boldsymbol{\varphi}}$ is the optimal value of $\boldsymbol{\varphi}$, which is found by solving the problem resulting from substituting (3.5.22) into (3.5.18), i.e.,

$$\hat{\boldsymbol{\varphi}}_i = \arg \min_{\boldsymbol{\varphi}} \left\| \mathbf{P}_h^\perp(\boldsymbol{\varphi})\mathbf{f}_i \right\|^2 \quad (3.5.23)$$

where

$$\mathbf{P}_h^\perp(\boldsymbol{\varphi}) = \mathbf{I} - \mathbf{P}_h(\boldsymbol{\varphi}) \quad (3.5.24)$$

$$\mathbf{P}_h(\boldsymbol{\varphi}) = \frac{\mathbf{h}(\boldsymbol{\varphi})\mathbf{h}^T(\boldsymbol{\varphi})}{\mathbf{h}^T(\boldsymbol{\varphi})\mathbf{h}(\boldsymbol{\varphi})} \quad (3.5.25)$$

and $\mathbf{P}_h(\boldsymbol{\varphi})$ is the projection matrix on the span of $\mathbf{h}(\boldsymbol{\varphi})$ and $\mathbf{P}_h^\perp(\boldsymbol{\varphi})$ is the complementary projection matrix of the latter.

To solve (3.5.23), iterative methods such as Gauss-Newton (GN) method or Levenberg-Marquardt (LM) method could be used. The estimate is updated iteratively according to

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

$$\hat{\varphi}^{k+1} = \hat{\varphi}^k + \epsilon \mathbf{d}^k \quad (3.5.26)$$

where ϵ is the step size in the search direction specified by \mathbf{d}^k and the superscript (k) denotes the iteration index. Note that the SN's index is dropped for notational convenience. On the other hand, \mathbf{d}^k is found by solving

$$\left((\mathbf{J}^k)^T \mathbf{J}^k + \epsilon \mathbf{I} \right) \mathbf{d}^k = \mathbf{g}^k \quad (3.5.27)$$

where $\epsilon = 0$ in case of GN and otherwise for LM and \mathbf{J}^k is the Jacobian matrix whose columns are defined as

$$\begin{aligned} [\mathbf{J}^k]_j &= -\frac{1}{\mathbf{h}^T(\varphi^k)\mathbf{h}(\varphi^k)} \left[\left(\mathbf{P}_h^\perp(\varphi^k) \frac{\partial \mathbf{h}(\varphi^k)}{\partial \varphi_j} \mathbf{h}^T(\varphi^k) \right) + \left(\mathbf{P}_h^\perp(\varphi^k) \frac{\partial \mathbf{h}(\varphi^k)}{\partial \varphi_j} \mathbf{h}^T(\varphi^k) \right)^T \right] \mathbf{f} \\ &\approx -\frac{\mathbf{P}_h^\perp(\varphi^k)}{\mathbf{h}^T(\varphi^k)\mathbf{h}(\varphi^k)} \frac{\partial \mathbf{h}(\varphi^k)}{\partial \varphi_j} \mathbf{h}^T(\varphi^k) \mathbf{f} \end{aligned} \quad (3.5.28)$$

where the subscript (j) in $[\mathbf{J}^k]_j$ refers to the j th column and in φ_j it refers to the j th element. The last line above follows if the residuals are small [126]. The analytical expressions of the derivatives of $\mathbf{h}(\varphi)$ are given in Appendix B.1. The gradient is computed as

$$\mathbf{g}^k = (\mathbf{J}^k)^T \mathbf{p}^k \quad (3.5.29)$$

and

$$\mathbf{p}^k = \mathbf{P}_h^\perp(\varphi^k) \mathbf{f}. \quad (3.5.30)$$

However, any gradient-based approach requires an initialization. Fortunately, (3.5.23) is *asymptotically convex*, i.e., as the number of samples goes to infinity the cost function becomes asymptotic. The following lemma provides the proof for the latter statement.

Lemma 3.1. *If the Jacobian matrix defined in (3.5.28) is full rank for all \mathbf{x} and φ , then the VP method (3.5.23) is asymptotically convex, i.e., as the number of measurements goes to infinity, problem (3.5.23) becomes convex.*

Proof. See Appendix B.4. □

This fact, theoretically, alleviates the need to find a good initial guess since

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

any will suffice, which in turn reduces the computational complexity of the local estimation problem.

3.5.2.2 Grid Search

The gradient approach requires matrix inversion, function evaluation, and derivative evaluation as well. Those operation are computationally demanding in the context of WSNs. Hence, we suggest the grid search (GS) method as an alternative when the SNs have low computational capacity. Although the GS might seem, at first glance, to require significant computation, we will show it to be otherwise.

Recall that φ is a two-dimensional parameter. Then φ can discretized into a two-dimensional grid. Let the parameter space be divided into K grids, then the grid values are the set $\{\varphi^k\}_{k=0}^{K-1}$ and the corresponding transfer vector is $\mathbf{h}_k = \mathbf{h}(\varphi^k)$. Thus, the cost function in problem (3.5.23) can be readily written as

$$\begin{aligned}
 \|\mathbf{P}_h^\perp(\varphi_k)\mathbf{f}\|^2 &= \mathbf{f}^T (\mathbf{P}_h^\perp(\varphi_k))^T \mathbf{P}_h^\perp(\varphi_k)\mathbf{f} \\
 &= \mathbf{f}^T \mathbf{P}_h^\perp(\varphi_k)\mathbf{f} \\
 &= \mathbf{f}^T \mathbf{f} - \mathbf{f}^T \frac{\mathbf{h}_k \mathbf{h}_k^T}{\mathbf{h}_k^T \mathbf{h}_k} \mathbf{f} \\
 &= \|\mathbf{f}\|^2 - \left\| \frac{\mathbf{h}_k^T \mathbf{f}}{\mathbf{h}_k} \right\|^2
 \end{aligned} \tag{3.5.31}$$

where the third equality follows from the symmetric and idempotent properties of \mathbf{P}_h^\perp . Using the above, the optimization problem (3.5.23) can be reformulated as

$$\hat{k} = \arg \max_k \frac{\|\mathbf{h}_k^T \mathbf{f}\|}{\|\mathbf{h}_k\|} \tag{3.5.32}$$

where the optimization variable here is the grid index. Given the solution to the above, then we have

$$\hat{\varphi} = \varphi^{\hat{k}} \tag{3.5.33}$$

$$\hat{a} = \frac{\mathbf{h}_{\hat{k}}^T \mathbf{f}}{\|\mathbf{h}_{\hat{k}}\|^2}. \tag{3.5.34}$$

The optimization problem (3.5.32) is merely finding the maximum dot product

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

between the normalized columns $\mathbf{h}_k / \|\mathbf{h}_k\|$ and the measurement vector. Hence, the implementation of (3.5.32) is simply a correlator of the received signal with the stored normalized vectors. At first glance, the storage requirement appears to be large if a good local estimates are desired, which is not the case indeed. Recall that $\boldsymbol{\varphi} = (b, n_\tau)^T$, where n_τ is the entry sample time and also it is the time shift of the time vector $\mathbf{h}(\boldsymbol{\varphi})$. Thus, we only need to store $\mathbf{h}(b, 0)$ and shift it according to n_τ . Furthermore, we can send \hat{k} (instead $\hat{\varphi}$) to the FC since the former is an integer value representing an index. Hence, it is expected to be represented by a small number of bits, which consequently requires less bandwidth.

3.5.3 Global Estimation Techniques

Here global estimation techniques are investigated. We start by using NLS to solve the localization problem. Then a grid search approach is formulated as an alternative to circumvent the need for a good initial guess. Finally, the localization problem is cast as a hyperspherical estimation form.

3.5.3.1 Gradient-based Search

Upon the reception of the intermediate estimates at the FC, it starts estimating the desired parameters ($\boldsymbol{\theta}$). This global estimation can be regarded as a fusion process of the local estimates to produce the desired result. The manner in which the fusion is carried out is not particularly specified by (3.3.4), as it is the case for local estimation. Hence, the fusion process can be approached from different angles. Next we present two methods for intermediate estimates fusion (or global estimation): Nonlinear least squares estimation, and the robust hyper-spherical intersection method.

Using the received intermediate estimates, we formulate the following nonlinear least squares estimation problem

$$(\hat{\mathbf{x}}_t, \hat{\kappa}) = \arg \min_{\mathbf{x}_t, \kappa} \sum_{i=0}^{M-1} \left(\hat{b}_i - \frac{\|\mathbf{x}_i - \mathbf{x}_t\|}{2\sqrt{\kappa T_s}} \right)^2. \quad (3.5.35)$$

The above can be solved by the VP approach via defining $\eta = 1/\sqrt{\kappa}$ as the linear parameter. In this case, the optimization problem to be solved is

$$\hat{\mathbf{x}}_t = \arg \min_{\mathbf{x}_t} \left\| \mathbf{P}_r^\perp(\mathbf{x}_t) \mathbf{b} \right\|^2 \quad (3.5.36)$$

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

where $\mathbf{b} = (\hat{b}_0, \dots, \hat{b}_{M-1})^T$ (with slight abuse of notation) and the complementary projection matrix is

$$\mathbf{P}_r^\perp(\mathbf{x}_t) = \mathbf{I} - \frac{\mathbf{r}(\mathbf{x}_t)\mathbf{r}^T(\mathbf{x}_t)}{\mathbf{r}^T(\mathbf{x}_t)\mathbf{r}(\mathbf{x}_t)} \quad (3.5.37)$$

and $\mathbf{r}(\mathbf{x}_t) = (\|\mathbf{x}_0 - \mathbf{x}_t\|, \dots, \|\mathbf{x}_{M-1} - \mathbf{x}_t\|)^T$. Then, the diffusivity estimate follows as

$$\hat{\kappa} = \frac{\mathbf{r}^T(\hat{\mathbf{x}}_t)\mathbf{r}(\hat{\mathbf{x}}_t)}{4T_s\mathbf{r}^T(\hat{\mathbf{x}}_t)\mathbf{b}}. \quad (3.5.38)$$

The cost function in (3.5.35) is plotted in Fig. 3.5.2. Obviously, the problem is not convex and requires a good initialization. For this purpose, we suggest using the centroid of the locations of the active SNs, i.e., the average of the detecting SNs locations.

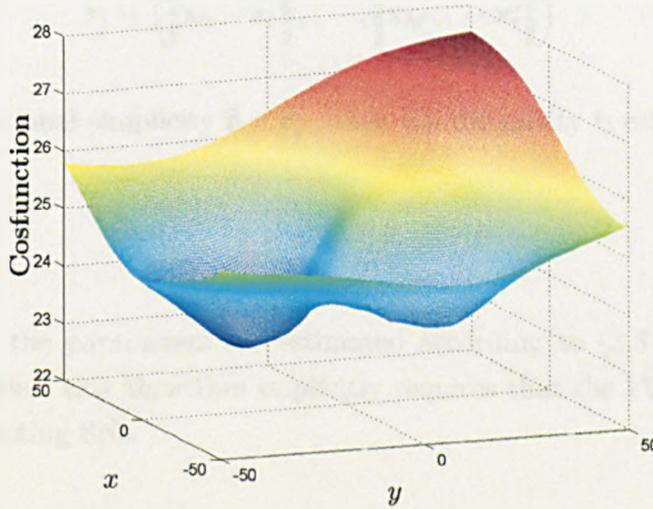


Figure 3.5.2: Global Cost Function.

3.5.3.2 Grid Search

To alleviate the need of a good initial guess and complicated function evaluations, (as before) a grid search (GS) is suggested as an alternative (a similar approach is adopted in [127]). Applying the procedure presented in (3.5.31) on (3.5.35) we have

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

$$\hat{\mathbf{x}}_t = \arg \max_{\mathbf{x}_t} \frac{\|\mathbf{r}^T(\mathbf{x}_t)\mathbf{b}\|}{\|\mathbf{r}(\mathbf{x}_t)\|}. \quad (3.5.39)$$

To implement the GS, first the sensing field (\mathcal{F}) is divided into M grid points each with center \mathbf{x}_t^j , which are the potential target locations. Let

$$\mathbf{r}_j = \left(\|\mathbf{x}_0 - \mathbf{x}_t^j\|, \dots, \|\mathbf{x}_{M-1} - \mathbf{x}_t^j\| \right)^T \quad (3.5.40)$$

be the distance between a SN location \mathbf{x}_i and the grid point \mathbf{x}_t^j . Then, similar to the local processing case, problem (3.5.19) reduces to

$$\hat{j} = \arg \max_j \frac{\|\mathbf{r}_j^T \mathbf{b}\|}{\|\mathbf{r}_j\|}. \quad (3.5.41)$$

Consequently, we have $\hat{\mathbf{x}}_t = \mathbf{x}_t^{\hat{j}}$ and it follows that

$$\mathbf{r}_{\hat{j}} = \left(\|\mathbf{x}_0 - \mathbf{x}_t^{\hat{j}}\|, \dots, \|\mathbf{x}_{M-1} - \mathbf{x}_t^{\hat{j}}\| \right)^T. \quad (3.5.42)$$

Let for notational simplicity $\hat{\mathbf{r}} = \mathbf{r}_{\hat{j}}$, then the diffusivity is estimated as

$$\hat{\kappa} = 4T_s \frac{\|\hat{\mathbf{r}}\|^4}{(\mathbf{b}^T \hat{\mathbf{r}})^2}. \quad (3.5.43)$$

The rest of the parameters are estimated according to (3.5.21) and (3.5.20). Note, however, that this algorithm implicitly requires that the FC knows the locations of the detecting SNs.

3.5.3.3 Hyper-Spherical Intersection (HSI) Method

However, the main drawback of the iterative NLSE approach is its computational load that requires cost function evaluation and matrix inversion. To overcome this hurdle several closed form approaches are suggested such as the HSI method [123], which performs localization by solving the hyper-spherical equations. We first investigate using the conventional HSI method then develop a robust approach.

The global localization problem (3.5.35) is formulated into the following HSI equations

3.5. SEGMENTED DISTRIBUTED LOCALIZATION AND ESTIMATION

$$\|\mathbf{x}_t - \mathbf{c}_{ij}\|^2 = \rho_{ij}^2 \quad (3.5.44)$$

where

$$\mathbf{c}_{ij} = \frac{1}{1 - K_{ij}^2} (\mathbf{x}_i - K_{ij}^2 \mathbf{x}_j) \quad (3.5.45)$$

are the center points and the corresponding radii are

$$\rho_{ij} = \frac{1}{1 - K_{ij}^2} \|\mathbf{x}_i - \mathbf{x}_j\| \quad (3.5.46)$$

with

$$K_{ij} = \frac{\hat{b}_i}{\hat{b}_j} \quad (3.5.47)$$

the ratio between the i th and j th SNs (b) intermediate estimates. Note, however, if the estimates are perfect, then the previous ratios would be

$$K_{ij} = \frac{\|\mathbf{x}_i - \mathbf{x}_t\|}{\|\mathbf{x}_j - \mathbf{x}_t\|}. \quad (3.5.48)$$

It is well known that the nonlinear least square problem (3.5.44) can be converted into a linear least square (LLS) [123] by taking an arbitrary SN as a reference node, say SN 0, and then forming pairs of hypersphere equations to acquire the overdetermined linear system

$$\underbrace{2(\mathbf{c}_{i0} - \mathbf{c}_{j0})^T}_{\mathbf{u}_i^T} \mathbf{x}_t = \underbrace{(\mathbf{c}_{i0}^2 - \rho_{i0}^2) - (\mathbf{c}_{j0}^2 - \rho_{j0}^2)}_{\zeta_i} \quad (3.5.49)$$

for every pair of active SNs. Compiling all the values in matrix form we have

$$\mathbf{U} \mathbf{x}_t = \boldsymbol{\zeta} \quad (3.5.50)$$

where $\mathbf{U} = (\mathbf{u}_0^T, \dots, \mathbf{u}_{L-1}^T)$, $\boldsymbol{\zeta} = (\zeta_0^T, \dots, \zeta_{L-1}^T)^T$, and L is the number of equations. Then, the target location is found by solving

$$\min_{\mathbf{x}_t} \|\mathbf{U} \mathbf{x}_t - \boldsymbol{\zeta}\| \quad (3.5.51)$$

3.6. SIMULATION RESULTS AND DISCUSSION

which is a linear least squares (LLS) problem that has the following closed form solution

$$\hat{\mathbf{x}}_t = (\mathbf{U}^T \mathbf{U})^{-1} \mathbf{U}^T \boldsymbol{\zeta} \quad (3.5.52)$$

given that \mathbf{U} is full rank. This condition is satisfied when having three or more detecting SNs.

Although the LLS is generally attractive, it is known to perform poorly in the presence of outliers, which are possible in the DDLE algorithm since a false alarm detection leads to erroneous intermediate parameters. Therefore, we propose solving (3.5.50) using the robust and convex Huber cost function [128] instead of the conventional l_2 -norm. In other words, we localize the target by solving

$$\hat{\mathbf{x}}_t = \arg \min_{\mathbf{x}_t \in \mathcal{F}} H(\mathbf{U} \mathbf{x}_t - \boldsymbol{\zeta}) \quad (3.5.53)$$

where \mathcal{F} is the sensing field and $H(\cdot)$ is the Huber function defined here as

$$H(\mathbf{u}) = \sum_{n=0}^{N-1} H_n(u_n), \quad \mathbf{u} \in \mathbb{R}^n \quad (3.5.54)$$

$$H_n(u_n) = \begin{cases} u_n^2 & , |u_n| \leq \epsilon \\ 2|u_n| - 1 & , |u_n| > \epsilon. \end{cases} \quad (3.5.55)$$

For small errors, the Huber function behaves like the square function, whereas for large errors it behaves as the l_1 -norm which is robust against outliers. Note that (3.5.55) reduces to LS when $\epsilon \rightarrow \infty$. Here, we choose $\epsilon = 1$ for simplicity.

Now, (3.5.53) can be solved iteratively using convex optimization techniques [128]. However a solution is guaranteed since (3.5.55) is a convex function. Therefore, localization robustness is traded-off with more computation complexity. The rest of the parameters are estimated according to (3.5.38), (3.5.21), and (3.5.20).

3.6 Simulation Results and Discussion

In this section we first present numerical simulation of the diffusion signal caused by a stationary target. Then the CRLB for both the target's and the medium's parameters are computed for different network setups. Finally, the mean square estimation

3.6. SIMULATION RESULTS AND DISCUSSION

(MSE) of the SDLE variants are compared together and with the CRLB.

3.6.1 Simulation Setup

Wireless Sensor Network

A WSN is randomly deployed in a sensing field, \mathcal{F} , of square shape with side length of $L = 100$ meters in the xy -plane. The deployment is modeled as a PPP $\Phi = \{\mathbf{X}_i\}$ with intensity λ sensor node per unit area. The network is simulated for 1000 Monte Carlo iterations where a new point process realization is generated in each iteration. One such realization is shown in Fig. 3.6.1.

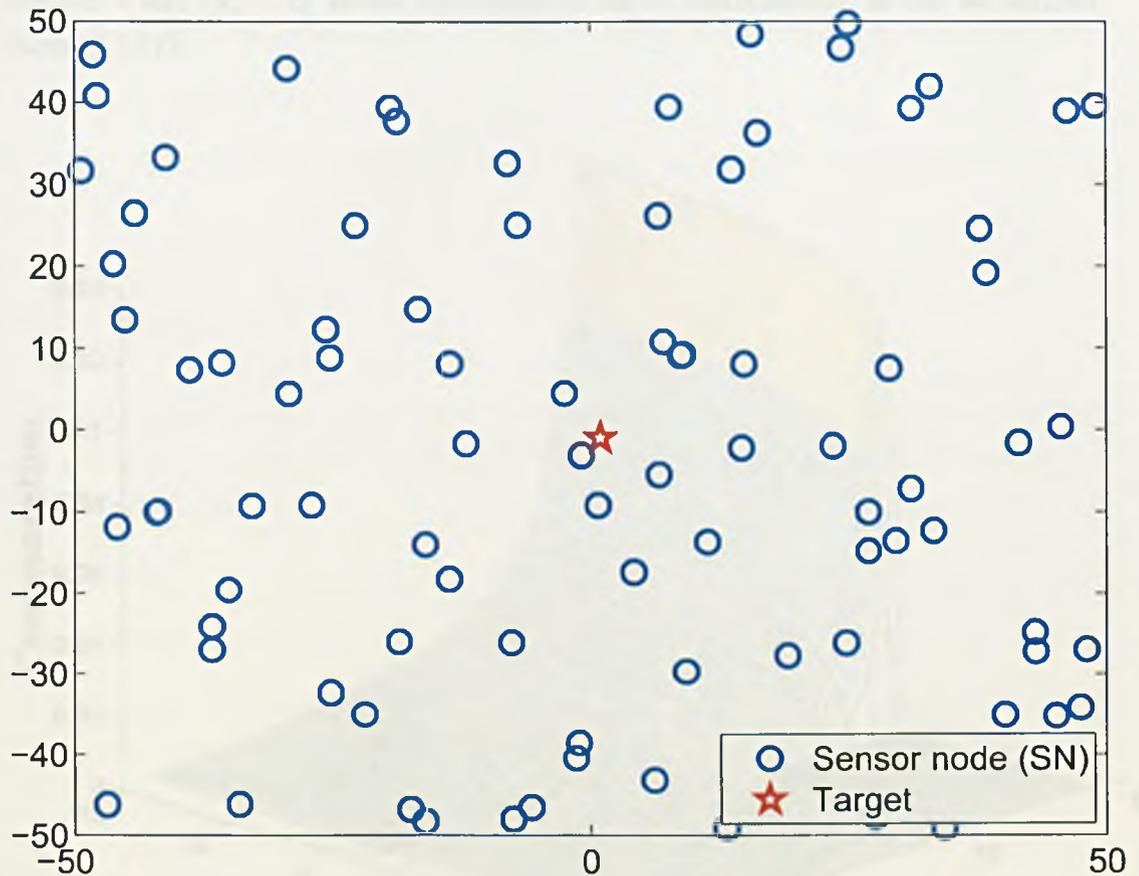


Figure 3.6.1: Wireless sensor network realization in sensing field \mathcal{F} , with target located at $(1, -1)$ for $\lambda = 10^{-2}$.

A stationary target is located at the arbitrary position $(1, -1, 0)$, at ground level for simplicity, entering the sensing field at time $\tau = 3.7$ sec. The target releases material with rate $\mu = 0.5\text{Kg/sec}$ into a medium with diffusivity of $\kappa = 40\text{m}^2/\text{sec}$. The arbitrary parameters chosen are

3.6. SIMULATION RESULTS AND DISCUSSION

$$\boldsymbol{\theta} = \begin{pmatrix} \mu \\ \mathbf{x} \\ \kappa \\ \tau \end{pmatrix} = \begin{pmatrix} 500 \\ 1 \\ -1 \\ 0 \\ 40 \\ 3.7 \end{pmatrix}. \quad (3.6.1)$$

The target causes a spatio-temporal variation in the concentration across the sensing field as shown in Fig. 3.6.2. Obviously, the signal degrades as the distance from the target increases. In contrast, the signal increases as time progresses because the target is constantly releasing material into the medium. The concentration stabilizes after some time that depends on μ , κ , and $\|\mathbf{x}_t - \mathbf{x}\|$. Larger μ values and smaller κ and $\|\mathbf{x}_t - \mathbf{x}\|$ values contribute to faster stabilization, as can be verified from (3.2.11).

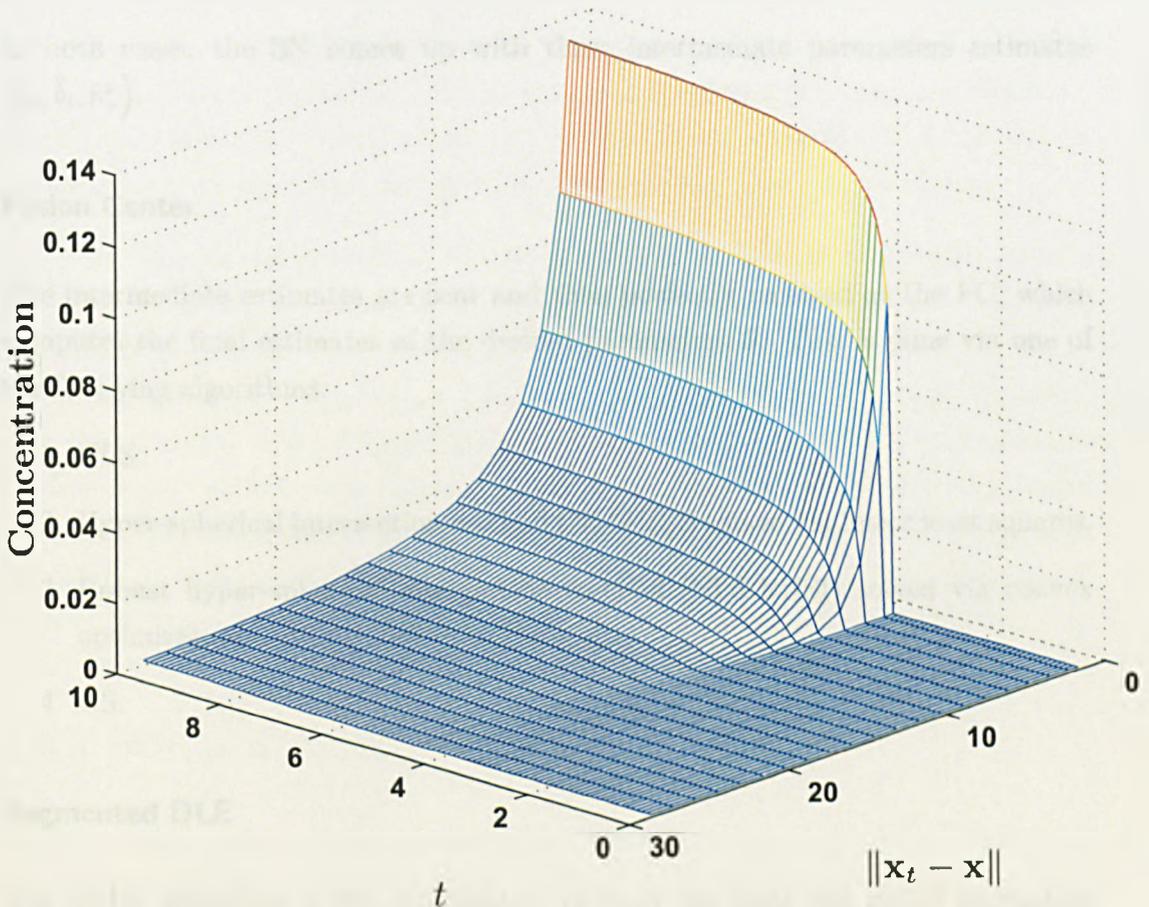


Figure 3.6.2: Concentration values plotted versus both time (t) and distance from target ($\|\mathbf{x}_t - \mathbf{x}\|$).

3.6. SIMULATION RESULTS AND DISCUSSION

Sensor Nodes

The SNs collect $N = 100$ samples corrupted by AWGN and with a known variance σ_s^2 that is set in our simulation to provide an average SNR of 20 dB. The SNR is defined as

$$\text{SNR}_s = \frac{1}{NL^2} \int_{\mathcal{F}} \sum_{n=n_r+1}^{N-1} s^2[n, \mathbf{x}] d\mathbf{x} \quad (3.6.2)$$

Upon collecting the samples, the SNs perform local detection via an energy detector with threshold γ . As explained earlier, local estimation is carried out if the prior distributed detection result is positive. The SNs locally estimate the intermediate parameters according to one of two algorithms:

1. Nonlinear least squares (NLS) estimation via the variable projection method.
2. Grid search (GS).

In both cases, the SN comes up with three intermediate parameters estimates $(\hat{a}_i, \hat{b}_i, \hat{n}_r^i)$.

Fusion Center

The intermediate estimates are sent and then perfectly received at the FC, which computes the final estimates of the desired parameters θ . This is done via one of the following algorithms:

1. NLS.
2. Hyper-spherical intersection method (NLS-HSI) solved via linear least squares.
3. Robust hyper-spherical intersection method (NLS-RHSI) solved via convex optimization.
4. GS.

Segmented DLE

The SDLE algorithm is the combination of both the local and global estimation approaches mentioned above. The algorithms used in the simulation are as follows:

The first three entries of the table are iterative algorithms and the fourth algorithm is the grid search. The last entry is the hybrid of the previous two.

3.6. SIMULATION RESULTS AND DISCUSSION

SDLE	Local Estimator	Global Estimator
NLS ²	NLS	NLS
NLS ² -HSI	NLS	NLS-HSI
NLS ² -RHSI	NLS	NLS-RHSI
GS ²	GS	GS
GS-NLS	GS	NLS

Table 3.1: Segmented distributed localization and estimation algorithms used in simulation.

3.6.2 CRLB

In the following numerical simulation, we investigate the behavior of the CRLB of the estimated parameters (θ) with respect to the local detection threshold (censoring threshold) γ and the network deployment density λ . We first set $\lambda = 3 \times 10^{-2}$ SN/m² and calculate the CRLB for θ for different sensing SNR (SNR_s) values as in Fig. 3.6.3. As expected, the CRLB improves as the SNR increases.

Similarly, setting $\lambda = 3 \times 10^{-2}$ and SNR_s = 20dB, we compute the CRLB for varying γ values. Fig. 3.6.4 illustrates the square root of the CRLB for θ given in (3.6.1). Except for the μ CRLB, changing the censoring threshold does not significantly affect the CRLB of the parameters. Hence, a small CRLB loss can be traded-off with a significant saving in communication power and bandwidth.

This time we set $\gamma \approx 190$, which corresponds to a local probability of false alarm to 10^{-7} , and λ is varied and the CRLB is observed as in Fig. 3.6.5. Clearly, the CRLB is inversely proportional to λ . This implies, as is also shown, that the CRLB asymptotically approaches zero as λ tends to infinity. However, it is clear that the reduction in the CRLB becomes small as λ increases. This behavior can be explained physically by referring to the fact that having a large number of SNs in a fixed area necessarily means that they are located close to each other. Consequently, SNs close to each other would provide similar information. Hence, the improvement in the CRLB would not be significant.

3.6.3 MSE of SDLE Algorithms

Here we investigate the MSE of the SDLE algorithms, (see Table 3.1) with respect to λ and γ . Firstly, we compare the MSE of the intermediate parameters' local estimation for the NLS and GS algorithms for different γ values in Fig. 3.6.7. The NLS algorithm is the optimal estimator for the case of Gaussian noise. Thus it is expected to have better performance in general. However, the GS algorithm achieves a similar performance when γ takes relatively high values. This is due to the fact

3.6. SIMULATION RESULTS AND DISCUSSION

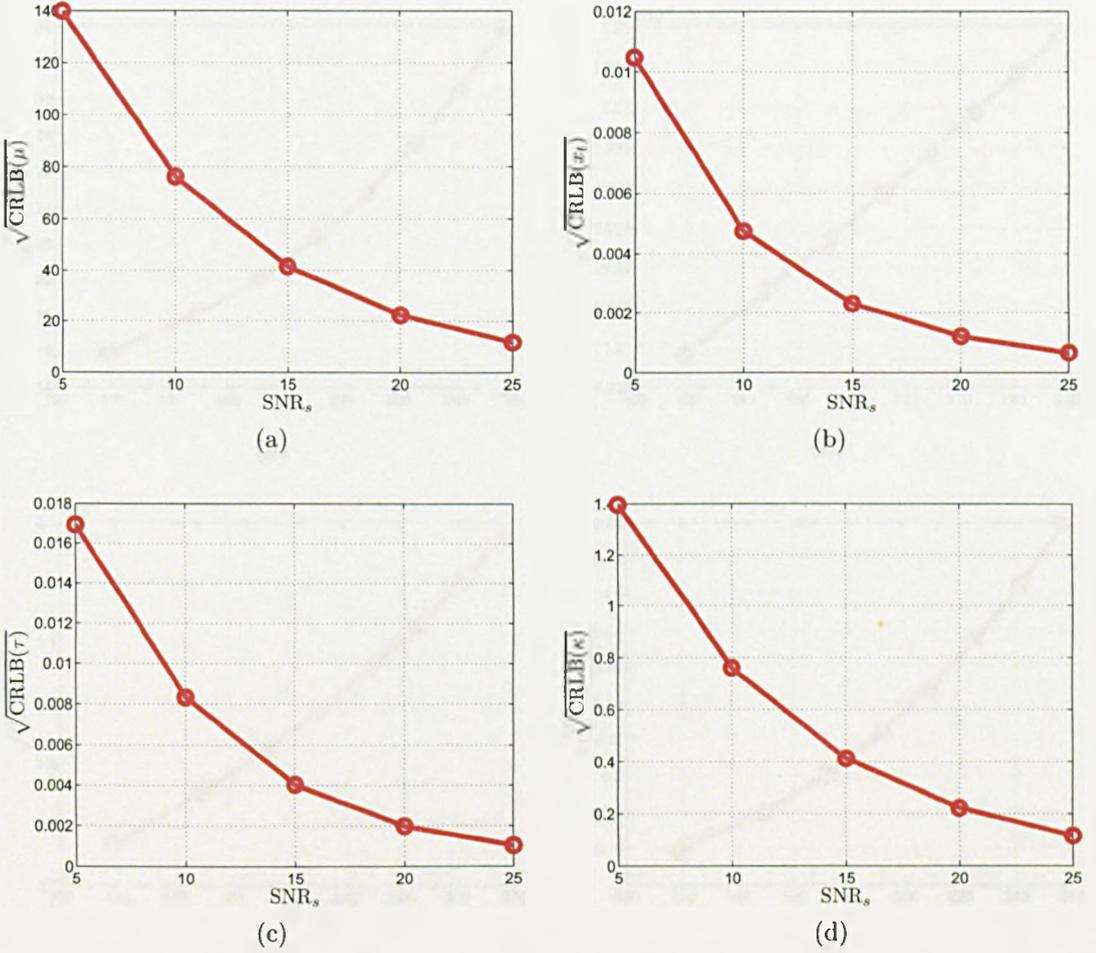


Figure 3.6.3: CRLB of θ given in (3.6.1) plotted against the sensing SNR (SNR_s) with $\gamma = 191$ corresponding to local probability of false alarm of 10^{-7} .

that higher γ results in the local detection of SNs with better information (closer to the target). Therefore the processed data will have a better SNR and consequently better estimation is attained. For example, Fig. 3.6.6 shows the effect of γ on the location and number of detecting SNs. If γ is set to be a high value, then the detecting SNs are concentrated around the target, in contrast to a relatively low value of it that causes a dispersion of the detecting SNs in the sensing field.

Next we show the MSE of the global estimation when γ is varied in Fig. 3.6.8. An obvious observation is that the MSE improves for all algorithms as γ increases for the same reason as mentioned above.

Starting with the iterative SDLE algorithms only, we note that the NLS² algorithm provides the best performance for all γ values, whereas NLS²-HSI gives the worst due to having erroneous intermediate estimates. This happens when a SN is located far away from the target and the measured signal is, consequently,

3.6. SIMULATION RESULTS AND DISCUSSION

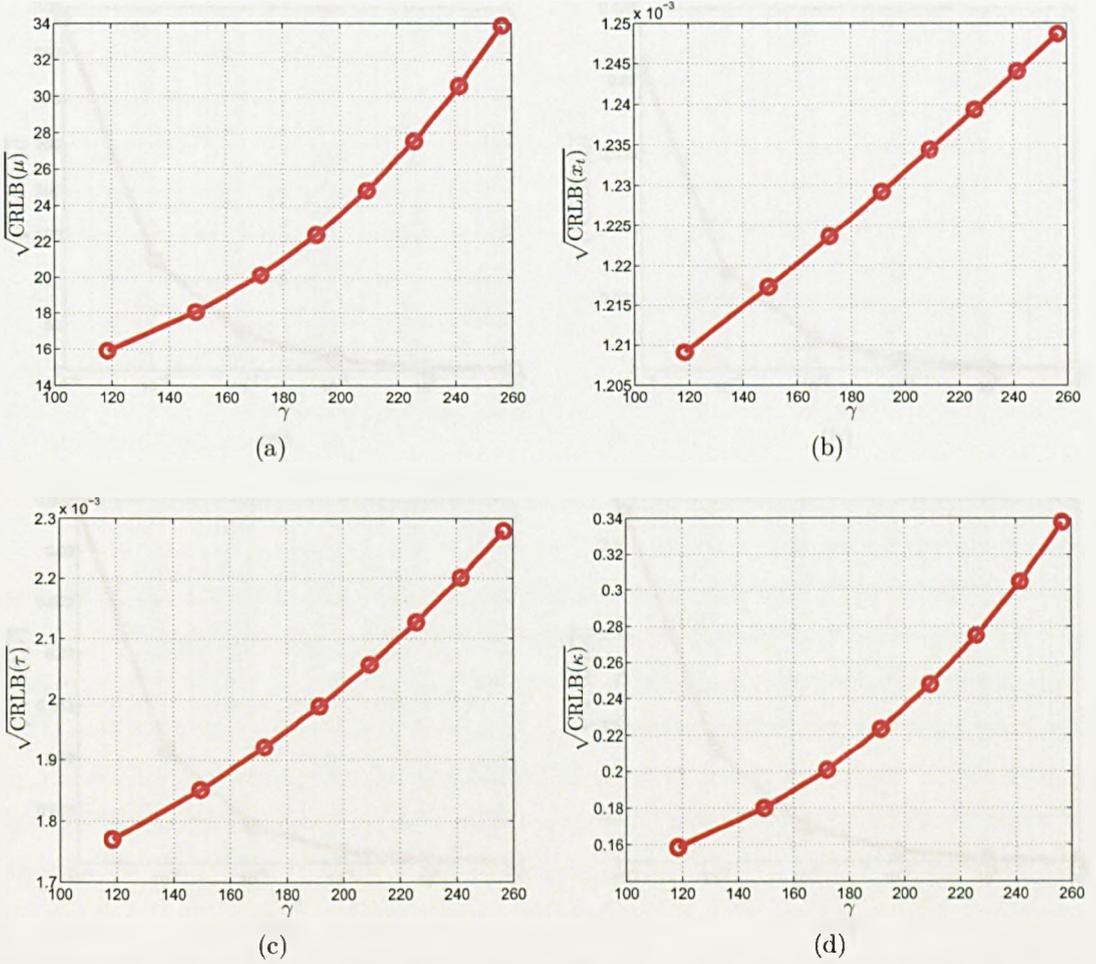


Figure 3.6.4: CRLB of θ for values given in (3.6.1) plotted against γ with $\lambda = 3 \times 10^{-2}$ SN/m².

dominated by noise. This leads to a *false alarm* type detection, also known as spurious detection [74]. Thus, the estimated parameters are the result of fitting the noise in the samples. In other words, those erroneous estimates act as outliers in the LLS localization, which is known to significantly degrade performance. On the other hand, the NLS²-RHSI performs better than the latter algorithm because it is robust to outliers. However, it does not reach the same performance level as NLS² since it inherits the properties of the hyper-spherical intersection algorithm that is known to be inferior to the NLS. It is shown that GS² catches up with the NLS as γ increases due to having a better SNR as stated before. The GS-NLS on the other hand, combines the optimal properties of the local NLS and the global properties of the GS. Hence it performs better than the rest. The previously discussed trends are apparent in all the estimated parameters. As for estimate of τ , two methods are used here to estimate n_τ^i that is used in (3.5.20) to estimate τ . It is clear that the GS performs much better than the NLS. This is due to the fact that the GS tries out many values and chooses the best one.

3.6. SIMULATION RESULTS AND DISCUSSION

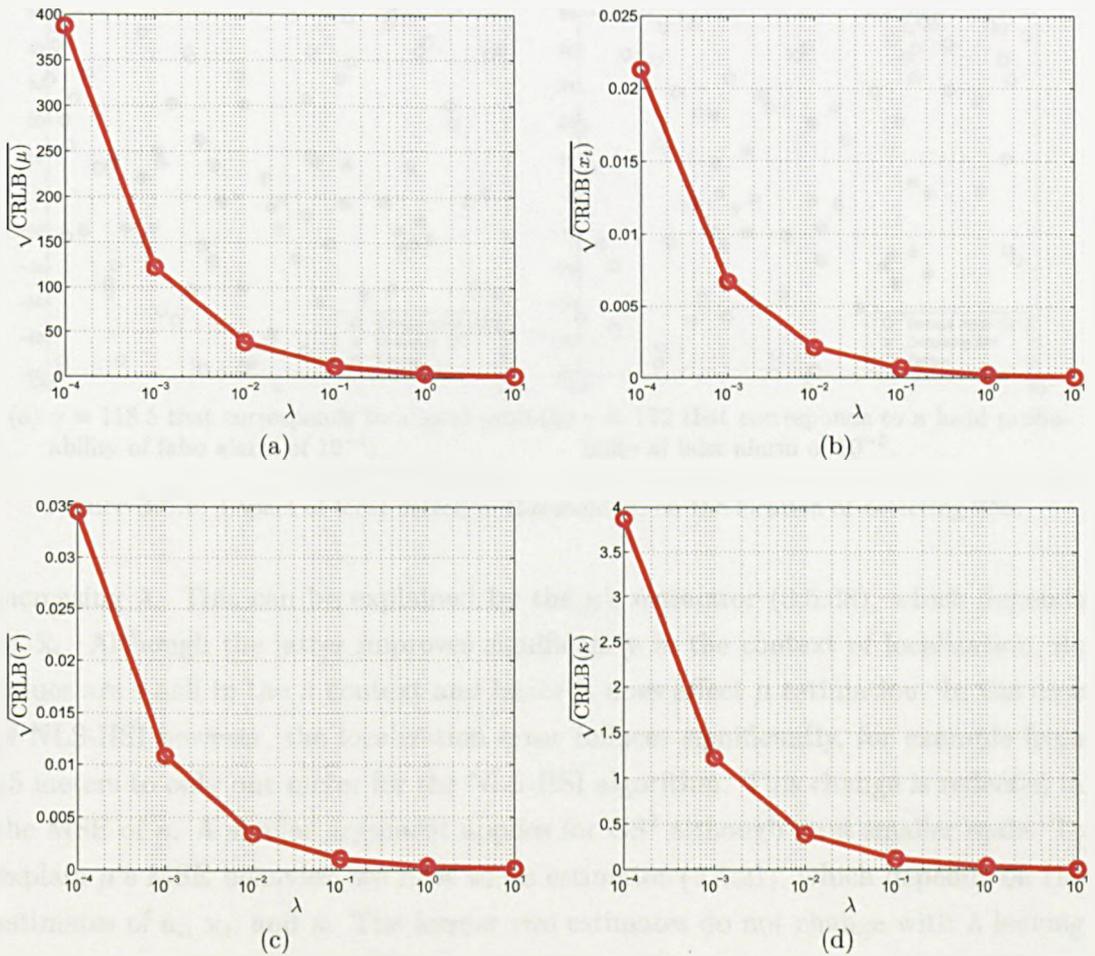
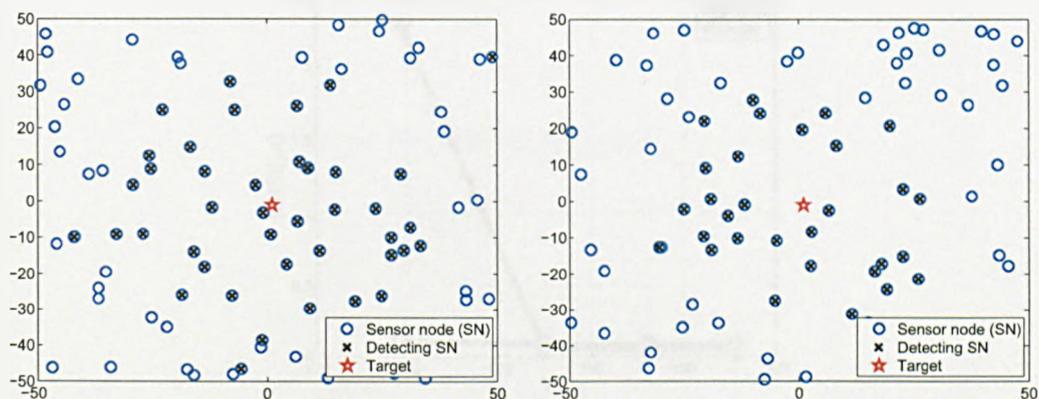


Figure 3.6.5: CRLB of θ given in (3.6.1) plotted against λ with $\gamma = 191$ corresponding to a local probability of false alarm of 10^{-7} .

Another impact of γ on SDLE is the number of detecting SNs. Fig. 3.6.9 shows the ratio of the average number of detecting SNs to the average number of deployed SNs as a function of γ . It can be seen that increasing γ will reduce the number of detecting SNs and hence save energy. For example, increasing γ from 118.5 to approximately 149 gives around a 14% reduction in the number of detecting SNs, whereas the loss of localization performance is negligible in the case of the NLS² algorithm.

Finally, the relationship between the MSE and λ is investigated in Fig. 3.6.10. The figures reveal an overall improvement in localization error for all the SDLE algorithms when increasing λ , which is intuitive since more information is available that in turn improves the estimation. In fact, GS² and NLS-MP achieve a localization error of 0.8 and 0.6 meters respectively at $\lambda = 0.1 \text{ SN}/m^2$, i.e., the previous performance is reached if one SN is deployed in every $100m \times 100m$ area, which is very reasonable in terms of the system cost. As for κ estimation, we note that the NLS², NLS-RHSI, and MP-NLS do not experience a significant change in the MSE with

3.7. CHAPTER SUMMARY



(a) $\gamma = 118.5$ that corresponds to a local prob- (b) $\gamma = 172$ that corresponds to a local prob-
 ability of false alarm of 10^{-1} . ability of false alarm of 10^{-5} .

Figure 3.6.6: Impact of local detection threshold, γ , on the location of detecting SNs.

increasing λ . This can be explained by the κ 's estimator (3.5.38), which depends on $\hat{\mathbf{x}}_t$. Although the latter improves significantly in the context of localization, its values are small in the κ context and hence it does affect μ estimation. In the case of NLS-HSI however, the localization error reduces significantly, for example from 15 meters to only one meter for the NLS-HSI algorithm. This change is reflected in the MSE of κ . A similar argument applies for GS² although on a smaller scale. To explain μ 's MSE behavior, we refer to its estimator (3.5.21), which depends on the estimates of a_i , \mathbf{x}_t , and κ . The former two estimates do not change with λ leaving $\hat{\kappa}$ to dominate the estimate. This explains the similarity between μ and κ MSEs.

3.7 Chapter Summary

In this chapter we have reviewed the derivation of the physical model of a diffusion source in a semi-infinite field. Then, we formulated a corresponding statistical model parametrized on the desired parameters vector $\boldsymbol{\theta} = (\mu, \mathbf{x}_t^T, \kappa, \tau)^T$. With this model at hand, we derived the maximum likelihood estimator (MLE) of $\boldsymbol{\theta}$ for a randomly deployed WSN with a SN intensity λ . The Cramer Rao lower bound (CRLB) for $\boldsymbol{\theta}$ estimation is found analytically using stochastic geometry tools. Numerical results of the CRLB show that it diminishes arbitrary with increasing λ . However, the CRLB is not significantly affected by varying the local detection (censoring) threshold γ .

We proposed the segmented distributed localization and estimation (SDLE) framework to serve as an efficient alternative to the resource expensive (although

3.7. CHAPTER SUMMARY

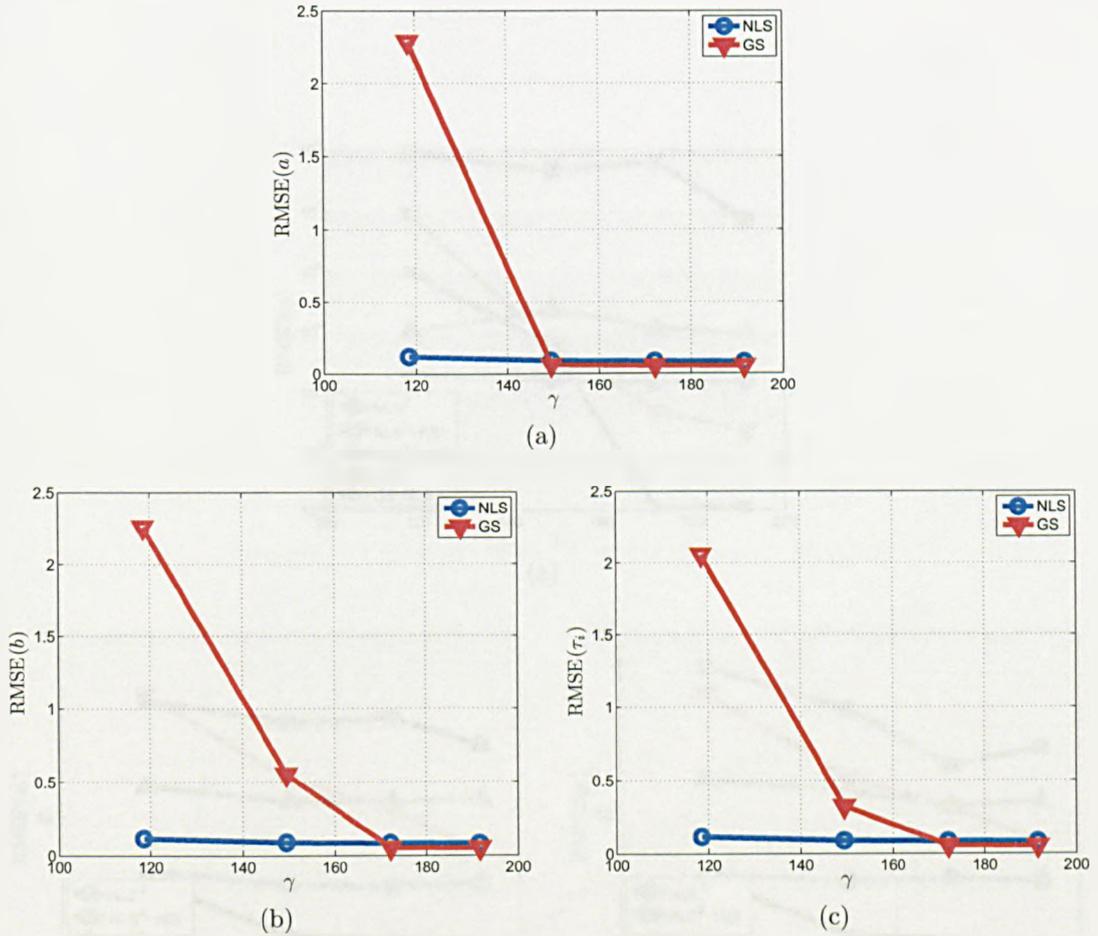


Figure 3.6.7: Local estimation RMSE of intermediate parameters, (a, b, n_r^t) , plotted against the local detection threshold γ , with $\lambda = 5 \times 10^{-2}$.

optimal) MLE. Two flavors are suggested: iterative algorithms and exhaustive search methods. The former is based on the nonlinear least squares formulation that is solved iteratively using algorithms like variable projection. The same problem is solved via grid search that turns out to be computationally simpler than the iterative methods. Similar to the CRLB behavior, the SDLE mean square error (MSE) is relatively insensitive to γ but reduces with increased λ as the simulation results shown. It is also shown that using a hybrid of iterative and sparse methods leads to a better MSE.

3.7. CHAPTER SUMMARY

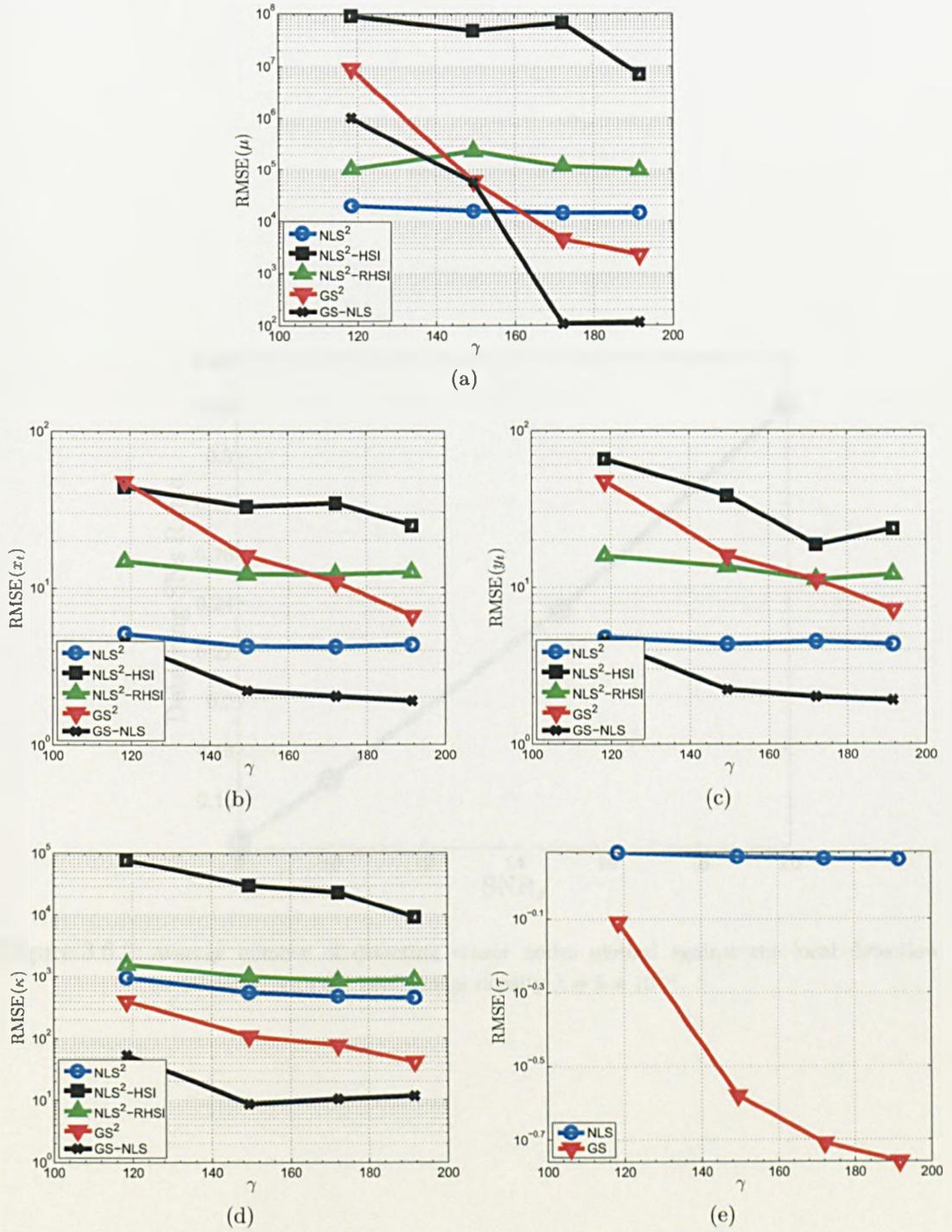


Figure 3.6.8: Global estimation RMSE of θ plotted against the local detection threshold γ , with $\lambda = 5 \times 10^{-2}$.

3.7. CHAPTER SUMMARY

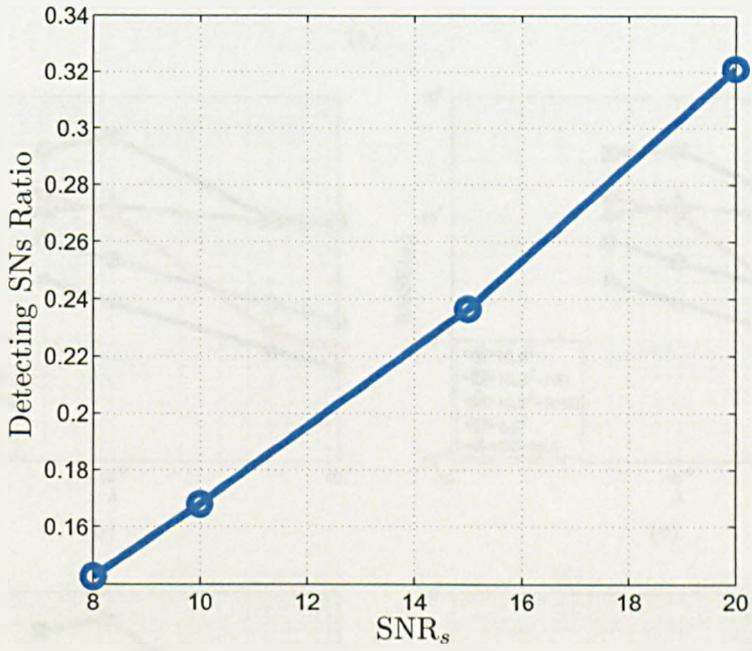


Figure 3.6.9: Average number of detecting sensor nodes plotted against the local detection threshold γ , with SN deployment density $\lambda = 5 \times 10^{-2}$.

3.7. CHAPTER SUMMARY

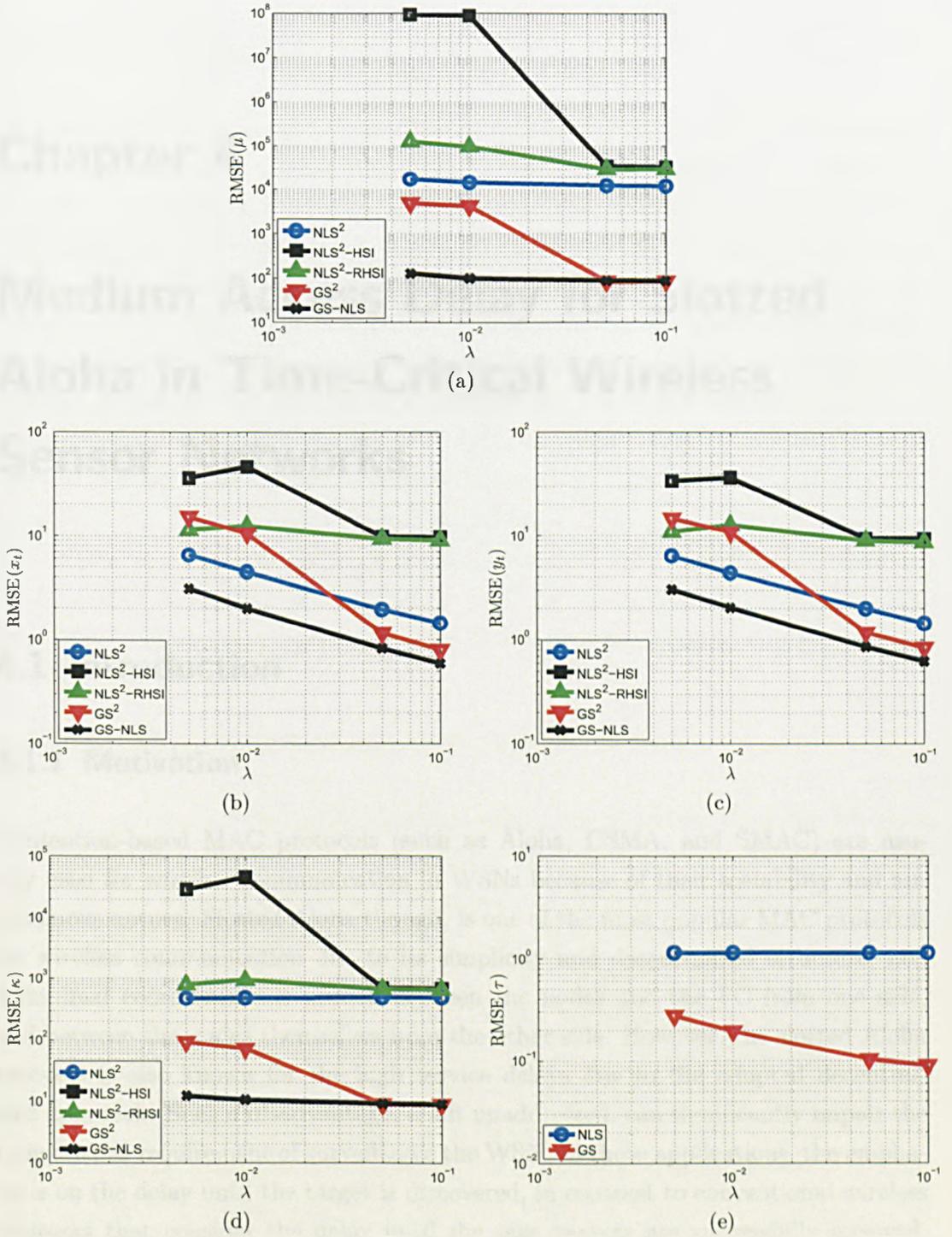


Figure 3.6.10: RMSE of estimators plotted against SN deployment density λ , with local detection threshold $\gamma = 172$ corresponding to local probability of false alarm of 10^{-5} .

Chapter 4

Medium Access Delay for Slotted Aloha in Time-Critical Wireless Sensor Networks

4.1 Introduction

4.1.1 Motivation

Contention-based MAC protocols (such as Aloha, CSMA, and SMAC) are usually used for wireless communication in WSNs because of their scalability and autonomous nature. Slotted Aloha though, is one of the most popular MAC protocols for wireless communication due to its simplicity and decentralized structure, i.e., a minimal coordination is needed between the nodes and the FC from one side, and between the nodes themselves from the other side. However, the slotted Aloha protocol is also known for the high service delays due to the adopted decentralized approach. Such a disadvantage, if left unaddressed, can significantly impair the time-critical requirement of surveillance the WSN. In those applications, the emphasis is on the delay until the target is discovered, in contrast to conventional wireless networks that consider the delay until the sent packets are successfully received. One proposed solution to reduce the medium access (contention) delay is to use multi-packet reception (MPR) via including several communication channels, such multiple frequencies, time slots, or codes. However, adopting MPR necessarily leads to an increase in hardware complexity, which is a crucial matter in WSN. Hence, it

4.1. INTRODUCTION

is vital to explore the performance improvement (in terms of delay statistics) when MPR is used in order to strike a trade-off between delay and complexity.

4.1.2 Related Work

Conventional contention delay distribution has been derived in [129] for a single node employing either slotted-Aloha or CSMA. The distributions were shown to intrinsically depend on the back-off policy used. The end-to-end delay distribution for a multi-hop WSN was investigated in [130], where a discrete-time Markov chain was used to model the system. The model encompasses both the queuing and the medium access delay. The authors in [131] considered the specific case of a decentralized detection WSN, in which the communication system was modeled as a continuous-time Markov process. The objective was to improve the detection error probability and the probability of delay violation, in contrast to traditional throughput and average delay metrics. Similarly, the work in [132] addressed the detection reliability of a WSN within a specific probabilistic delay bound through modeling the system of nodes and traffic as a spatio-temporal fluid. Stability and delay of MPR slotted Aloha was investigated in [133] for conventional wireless networks. Surveillance networks using MPR slotted Aloha were investigated in [134].

4.1.3 Chapter Proposal and Contributions

In this chapter, we statistically investigate the delay experienced by the FC until all the active SNs send their data successfully using slotted Aloha with and without MPR. To this end, we first derive the contention delay and statistically characterize its distribution in slotted Aloha. However, the latter distribution is very complicated, and so we resort to approximating it by a negative binomial distribution, which is the discrete counter part of the Gamma distribution. We use results from Chapter 2 to aid us in finding the mean and variance of the contention delay. Then, this information is used to tune the medium access probability of the MAC protocol to achieve a minimum delay. Thus, we address the case of MPR slotted Aloha in which the mean, variance, and distribution are analytically studied. We show (analytically and via simulations) that the delay statistics improves as the number of channels increases.

4.2. PROBLEM FORMULATION

4.1.4 Chapter Outline

In Section 4.2 the contention delay assumptions are presented and the problem is formulated. The contention delay in slotted Aloha is defined and its distribution is derived in Section 4.4. The MPR slotted Aloha contention delay is investigated in Section 4.4. Section 4.5 presents the simulation results and discussion. Finally, we give some conclusions in Section 4.6.

4.2 Problem Formulation

4.2.1 System Model

The system model used here is the same as the one used in Chapter 3 where the target is a diffusive source. This model is chosen in order to preserve continuity with the previous chapter and at the same time to show that the analysis in Chapter 2 also applies to diffusive sources. The SNs communicate with the FC over a shared channel. However we do not consider any imperfections here because, as discussed earlier, the SDLE framework enables low bandwidth transmission that in turn allows for powerful error correction codes. The data is sent only if the FC decides that a target is present and so transmission power is therefore saved. The MAC protocol used by the SNs is the p -persistent slotted-Aloha, in which every SN decides to transmit in the designated time slot with probability p_{ma} . The time slot is assumed to be of one time unit length for the sake of simplicity.

4.2.2 Contention Delay in DDLE

When the WSN attains a positive global detection decision about the target's presence, a distributed estimation stage is initiated. As explained before (see Fig. 4.2.1), the SNs estimate local intermediate parameters and then send them to the FC. The transmission occurs over a shared MAC. Since the data payload sent is relatively small, powerful error coding can be applied to protect the data. Therefore, we assume that the data are received perfectly at the FC. However, using slotted Aloha causes reception delay at the FC due to medium access collisions. In this chapter, we are concerned with investigating the total delay experienced by the FC until all the transmitting SNs successfully send their data through the shared channel to the FC.

In the context of the DDLE algorithm, we will be discussing part (III) high-

4.3. CONTENTION DELAY ANALYSIS IN SLOTTED ALOHA

lighted in Fig. 4.2.1. Our concern in this chapter is two folds. First, to statistically characterize the contention delay in WSN performing DDLE. Second, using the previous statistical description, we aim to choose the optimal medium access probability.

4.3 Contention Delay Analysis in Slotted Aloha

In this section, we define the delay from the FC point of view and approximate its distribution via moment matching with a negative binomial distribution. Finally, an insight into the medium access probability on the contention delay is provided.

Given the DDLE set up explained earlier, we expect to have Ω SNs ready to send their local estimates to the FC. This value is actually the offered traffic by the network. Recall that Ω is a RV given by

$$\Omega = \sum_{\mathbf{X}_i \in \Phi} I(\mathbf{X}_i) \quad (4.3.1)$$

as it was defined by (2.3.2). Note that we drop the subscript denoting hypothesis dependence since \mathcal{H}_1 is assumed. Moreover, recall that Ω is a Poisson RV, i.e.,

$$p_{\Omega}(\omega) = \text{Poi}(\bar{\Omega}) \quad (4.3.2)$$

where

$$\bar{\Omega} = \lambda \int_{\mathcal{F}} P_d(\mathbf{x}) d\mathbf{x} \quad (4.3.3)$$

and $P_d(\mathbf{x})$ is the detection probability at location \mathbf{x} .

Now Ω SNs will compete for the time slot. Each SN decides to send its data with probability p_{ma} . If the transmission is successful, the FC will acknowledge the sending SN and the latter will withdraw from the contention. If the transmission is not successful, the SN will try again in the next time slot with probability p_{ma} as well. The contention ends if there are Ω successful transmissions. Hence, the contention delay is the number of time slots required for Ω successful transmissions. This is in fact, the delay experienced by the FC to get all the required data, which is different from the delay experienced by the SN to send its packet.

4.3. CONTENTION DELAY ANALYSIS IN SLOTTED ALOHA

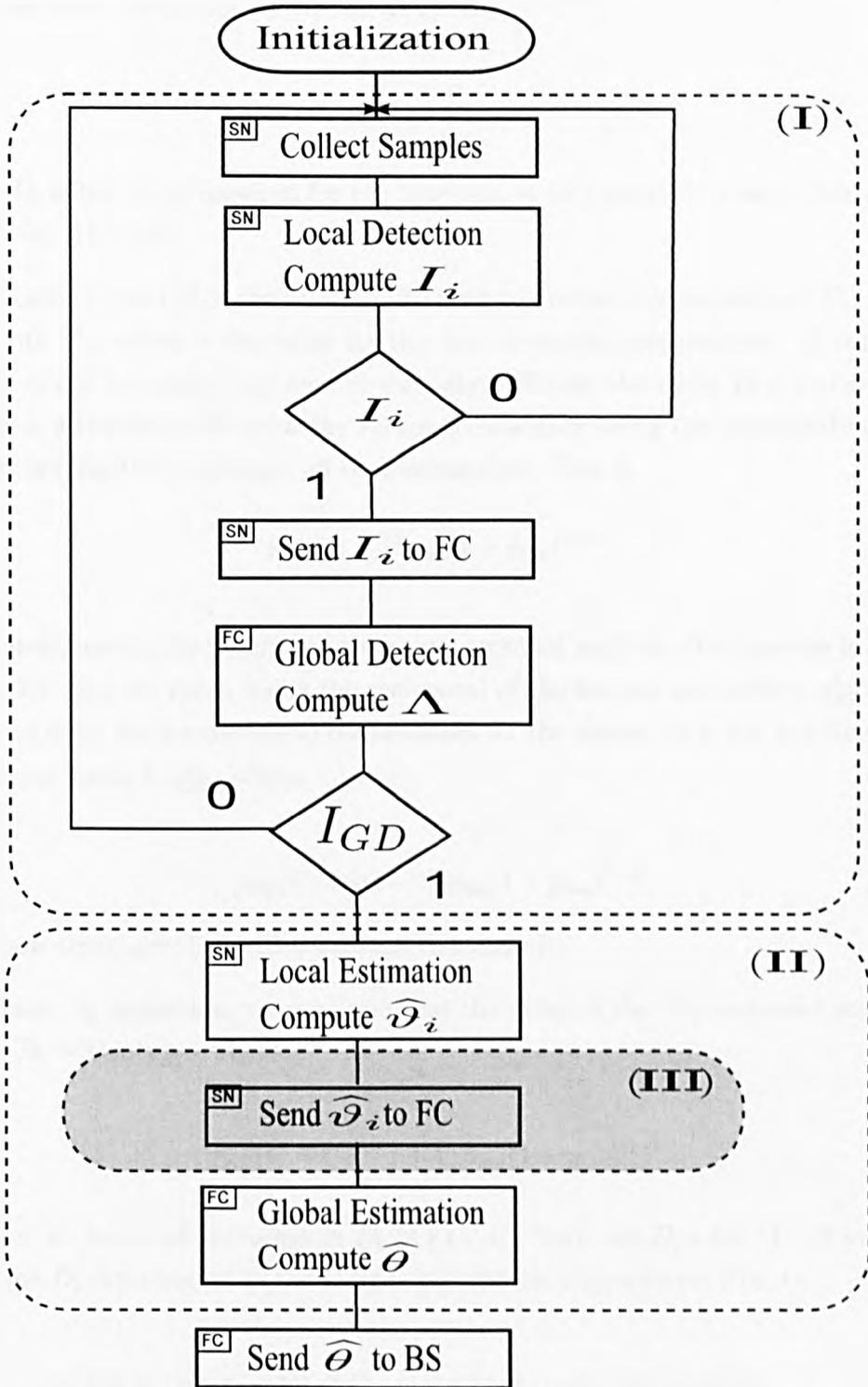


Figure 4.2.1: Distributed Detection, Localization and Estimation (DDLE) algorithm. In this chapter, section (III) is discussed.

4.3. CONTENTION DELAY ANALYSIS IN SLOTTED ALOHA

4.3.1 Exact Distribution

The contention delay can be defined as follows:

$$D_c = \sum_{i=1}^{\Omega} D_i \quad (4.3.4)$$

where D_i is the delay required for the time slot to be captured by any transmitting SN for the i th time.

In order to find D_c 's distribution, we need to discuss the statistics of D_i . Let us start with D_1 , which is the delay for the first successful transmission. In this case, Ω SNs would be competing for the time slot. Hence, the event of a transmission success is a Bernoulli RV with the success probability being the probability of only one SN transmitting amongst all contending SNs. This is

$$p_{ST}(1) = \Omega p_{ma}(1 - p_{ma})^{\Omega-1}. \quad (4.3.5)$$

Consequently, the number of time slots required until the first success is a Geometric RV with its mean being the reciprocal of the success probability, p_{ST}^1 . Similarly, the delay until a successful transmission for the second time slot is a Geometric RV having mean $1/p_{ST}^2$, where

$$p_{ST}(2) = (\Omega - 1) p_{ma}(1 - p_{ma})^{\Omega-2} \quad (4.3.6)$$

since now there are $\Omega - 1$ SNs wanting to transmit.

Hence, by induction, we conclude that the delay of the i th successful transmission is $D_i \sim \text{Geo}(p_{ST}^i)$ where

$$p_{ST}(i) = (\Omega - i + 1) p_{ma}(1 - p_{ma})^{\Omega-i}. \quad (4.3.7)$$

Now we turn our attention to D_c in (4.3.4). Since the D_i 's are in fact independent, the D_c distribution is the Ω -fold convolution of geometric RVs, i.e.,

$$p_{D_c}(d) = \text{Geo}(p_{ST}(1)) \otimes \text{Geo}(p_{ST}(2)) \otimes \cdots \otimes \text{Geo}(p_{ST}(\Omega)) \quad (4.3.8)$$

where \otimes is the convolution operator. Remember that Ω is also a RV.

4.3. CONTENTION DELAY ANALYSIS IN SLOTTED ALOHA

4.3.2 Negative Binomial Approximation

The above distribution is very complicated and does not have a closed form expression. Hence, we resort to approximation via moment matching. First, we estimate the first and second order moments and then match them with the first two statistics of the negative binomial distribution. We start by finding the mean delay:

$$\mathbb{E}[D_c] = \mathbb{E}_1 \left[\sum_{i=1}^{\Omega} D_i \right] \quad (4.3.9)$$

where the expectation is w.r.t to the joint distribution of Ω and the D_i 's under the \mathcal{H}_1 hypothesis. Fortunately, all the previous RVs are independent of each other. Thus we have

$$\begin{aligned} \mathbb{E}_1 \left[\sum_{i=1}^{\Omega} D_i \right] &= \mathbb{E}_1 \left[\sum_{i=1}^{\Omega} \mathbb{E}[D_i] \right] \\ &= \mathbb{E}_1 \left[\sum_{i=1}^{\Omega} \frac{1}{p_{st}(i)} \right] \end{aligned} \quad (4.3.10)$$

where $\mathbb{E}_1[\cdot]$ is the expectation under \mathcal{H}_1 and the expectation in the last equality is w.r.t Ω . On the other hand, the variance can be found using the total variance law as

$$\begin{aligned} \text{var}_1(D_c) &= \mathbb{E}_1[\text{var}(D_c|\Omega)] + \text{var}(\mathbb{E}_1[D_c|\Omega]) \\ &= \mathbb{E}_1 \left[\sum_{i=1}^{\Omega} \text{var}(D_i) \right] + \mathbb{E}_1 \left[\left(\sum_{i=1}^{\Omega} \mathbb{E}[D_i] - \mathbb{E}[D_c] \right)^2 \right] \end{aligned} \quad (4.3.11)$$

where for a given Ω , we have $\text{var}(D_i) = (1 - p_{st}(i))/p_{st}^2(i)$, which follows from D_i being a geometric RV.

Having found the first two moments, we approximate $p_{D_c}(d)$ by moment matching with a negative binomial distribution because it is flexible enough to take various forms. In fact, the negative binomial distribution is the discrete counterpart of the Gamma distribution. The negative binomial distribution (NB(r, q)) has parameters r and q , which are respectively the number of failed trials before the first success and the success probability. Thus

4.3. CONTENTION DELAY ANALYSIS IN SLOTTED ALOHA

$$\text{NB}(r, q) = \binom{k+r-1}{k} (1-q)^r q^k. \quad (4.3.12)$$

Using moment matching, the required distribution parameters are given by

$$q = \frac{\mathbb{E}_1[D_c]}{\text{var}_1(D_c)} \quad (4.3.13)$$

and

$$r = \frac{\mathbb{E}_1^2[D_c]}{\text{var}_1(D_c) - \mathbb{E}_1[D_c]}.$$

Now, the contention delay distribution can be approximately fitted to the following negative binomial distribution:

$$p_{D_c}(d) \approx \binom{d+r-1}{d} (1-q)^r q^d. \quad (4.3.14)$$

Having the contention delay distribution at hand enables the inspection of the real-time aspect of the DDLE system. In such a situation, we are interested in the event of satisfying or breaking the real-time deadline specified by the application. The real-time deadline miss probability captures the needed information about the system performance, which is defined as

$$P_{RM} = \mathbb{P}(D_c > TS) \quad (4.3.15)$$

where TS is the maximum required number of time slots to accommodate the communication operation. The latter probability resembles the outage probability used in communication systems. Obviously, P_{RM} is desired to be small, e.g., 1%. In other words, it is required that the real-time deadline is missed one percent of the time.

4.3.3 Medium Access Probability Impact

Apparently, there is no tractable relationship available between p_{ma} and the contention delay statistics. However, we will try to get an insight into the behavior of the contention delay w.r.t p_{ma} . We consider the mean of the contention delay (\bar{D}_c) given in (4.3.10). The latter can be written w.r.t p_{ma} as

4.4. CONTENTION DELAY ANALYSIS IN MPR SLOTTED ALOHA

$$\bar{D}_c = \mathbb{E}_1 \left[\sum_{i=1}^{\Omega} \frac{1}{(\Omega - i + 1) p_{ma} (1 - p_{ma})^{\Omega - i}} \right]. \quad (4.3.16)$$

Making the simple substitution ($j = i - 1$) enables further simplification to give

$$\bar{D}_c | \Omega = \frac{1}{p_{ma} (1 - p_{ma})^{\Omega - 1}} \sum_{j=0}^{\Omega - 1} \frac{1}{(\Omega - j) p_{ma} (1 - p_{ma})^{\Omega - j}} \quad (4.3.17)$$

where the average above is conditioned on Ω . Note that the summation multiplier above is nothing but the average delay in conventional Aloha scenarios, and so, we should expect a similar behavior. In other words, there exists an optimal p_{ma} value that minimizes the average contention delay. Nonetheless, the statistical expressions (4.3.10), (4.3.11), and (4.3.14) can be numerically evaluated to get the optimal p_{ma} value.

4.4 Contention Delay Analysis in MPR Slotted Aloha

As shown by the simulation results, a P_{RM} of 1% requires at best over 130 time slots, which is a prohibitive amount of delay. Thus, It is necessary to consider supporting technologies such as the MPR. In the MPR slotted Aloha, there exist a fixed number of channels (frequency, time slot, code) say L that any SN can use to transmit its data. When the SN decides to transmit, it chooses one of the L channels randomly and uniformly. To formally discuss the contention delay, we first need to define the channel choice variable at the SN, \mathbf{X}_i , as $C(\mathbf{X}_i)$, which can only take integer values in the set $\{1, \dots, L\}$. Let us also define the channel indicator at the SN \mathbf{X}_i as

$$I_l(\mathbf{X}_i) = \begin{cases} 1, & C(\mathbf{X}_i) = l \\ 0, & C(\mathbf{X}_i) \neq l \end{cases}. \quad (4.4.1)$$

Then the traffic offered at a the l th channel is

$$\Omega_l = \sum_{\mathbf{X}_i \in \Phi} I(\mathbf{X}_i) I_l(\mathbf{X}_i). \quad (4.4.2)$$

The traffic's distribution on the other hand is given by the following lemma.

4.4. CONTENTION DELAY ANALYSIS IN MPR SLOTTED ALOHA

Lemma 4.1. *If the traffic offered by the network is Poisson with mean $\bar{\Omega}$, then the traffic at the l th channel follows a Poisson distribution, i.e.,*

$$p_{\Omega_l}(\omega_l) = Poi\left(\frac{\bar{\Omega}}{L}\right). \quad (4.4.3)$$

Proof. It is easy to recognize that the traffic in (4.4.2) is a thinned version of the original offered traffic (4.3.1), because $I_l(\mathbf{X}_i)$ has a uniform probability of $1/L$. Hence, Ω_l is a Poisson RV with mean $\bar{\Omega}/L$. \square

The above lemma states that the traffic at any channel in an L MPR system is $1/L$ the original traffic, an intuitive result since the SNs choose the channel in a uniformly random manner.

Now the contention delay for the l th channel follows as

$$D_c^l = \sum_{i=1}^{\Omega_l} D_i^l \quad (4.4.4)$$

where D_i^l is the delay required for the time slot to be captured by any transmitting SN for the i th time in the l th channel. The D_c^l distribution can be found in a similar manner as described in Section 4.3. The final contention delay on the other hand, depends on the delay in all the channels. We define the contention delay in MPR slotted Aloha as the maximum of all the channel delays, i.e.,

$$D_c = \max(D_c^1, \dots, D_c^L). \quad (4.4.5)$$

The distribution of D_c now follows from order statistics as

$$p_{D_c}(d) = L \left(P_{D_c^l}(d) \right)^{L-1} p_{D_c^l}(d) \quad (4.4.6)$$

where $P_{D_c^l}(d)$ and $p_{D_c^l}(d)$ are the cumulative and probability distributions of the l th channel contention delay respectively. Consequently, the mean and variance can be computed from their basic definitions, i.e.,

$$\mathbb{E}[D_c] = \sum_{d=0}^{\infty} d p_{D_c^l}(d) \quad (4.4.7)$$

$$\text{var}[D_c] = \sum_{d=0}^{\infty} (d - \mathbb{E}[D_c])^2 p_{D_c^l}(d). \quad (4.4.8)$$

4.5. SIMULATION RESULTS AND DISCUSSION

4.5 Simulation Results and Discussion

We simulate a WSN in a square of $100\text{ m} \times 100\text{ m}$ with SNs deployed according to a uniform random distribution therein. The SNs are deployed with intensity $\lambda\text{ SN/m}^2$. Every SN takes 100 samples of a diffusive target characterized by the parameters values are chosen arbitrarily to be $\theta = (100, 0, 0, 40, 3.7)$, which represents the release rate, x -coordinate, y -coordinate, diffusivity, and entry time respectively. Note that the target is located at the origin for convenience. The sensing SNR is defined according to (3.6.2) and is set to 20 dB without loss of generality. The SNs perform local detection using the ED, and if the detection was positive it tries to transmit data to the FC over an ideal shared channel. All the SNs use slotted Aloha with the same medium access probability p_{ma} . We run the simulations for 10^4 Monte Carlo iterations.

Note that we assume knowledge of θ is available for simplicity. In practice, the approach 3.5.1.2 is suggested to find a lower bound on the performance.

4.5.1 Slotted Aloha

For the slotted Aloha without MPR, the contention delay mean, variance, and distributions behavior with p_{ma} is shown for different λ and γ values. Theoretical and simulated values are plotted. In Figures 4.5.1, 4.5.2, 4.5.3 and 4.5.4 the mean, variance (in which the theoretical delay mean and variance are given by equations (4.3.10) and (4.3.11) respectively), and TS (the required number of time slots to achieve a probability of real-time deadline miss of 0.1 given in (4.3.15)) are plotted against the medium access probability. Generally the figures indicate a clear similarity between the theoretical and simulated values. Clearly, increasing the λ will increase the offered traffic, which in turn decreases the optimum p_{ma} value. Changing γ , on the other hand, does not have a significant effect. We also note that the TS to achieve $P_{RM} = 0.1$ decreases when λ decreases due to the same reason stated earlier.

4.5.2 MPR Slotted Aloha

Fig. 4.5.5 depicts theoretical and simulated contention delay statistics plotted against the MPR channel number with p_{map} value set to the optimal value as shown in the figures above. Generally, the delay decreases as the number of channels increases, which is expected because the traffic per channel decreases.

4.5. SIMULATION RESULTS AND DISCUSSION

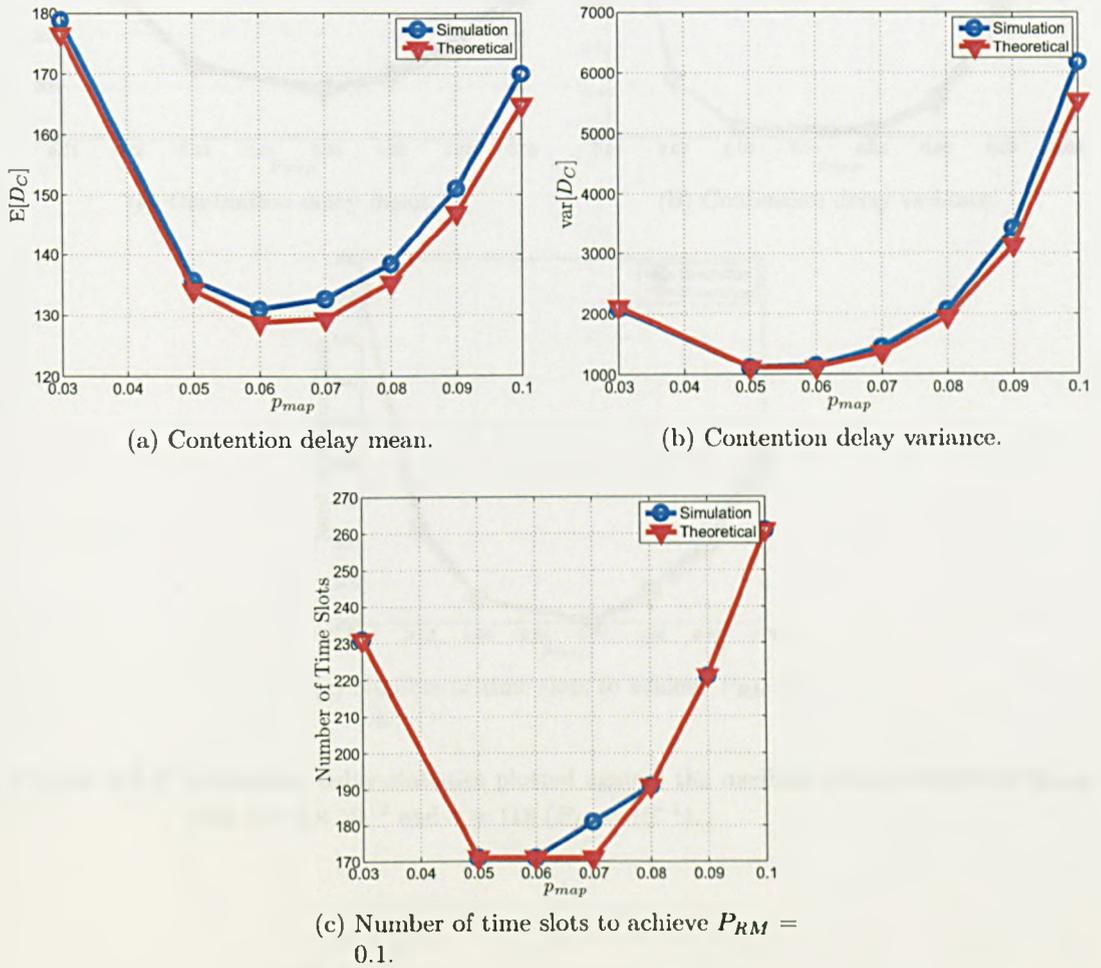


Figure 4.5.1: Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 1 \times 10^{-2}$ and $\gamma = 172$ ($P_{fa} = 10^{-5}$).

4.5. SIMULATION RESULTS AND DISCUSSION

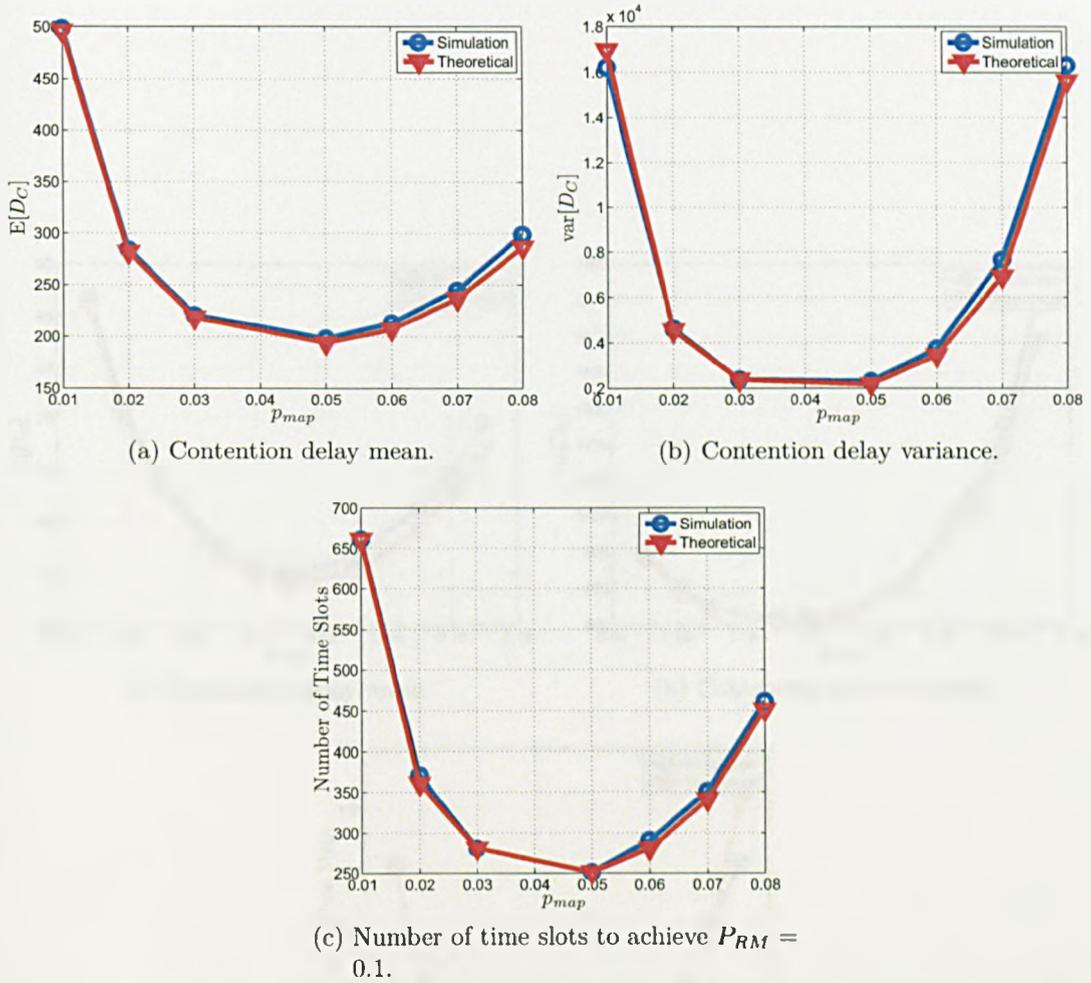


Figure 4.5.2: Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 1 \times 10^{-2}$ and $\gamma = 118$ ($P_{fa} = 10^{-1}$).

4.5. SIMULATION RESULTS AND DISCUSSION

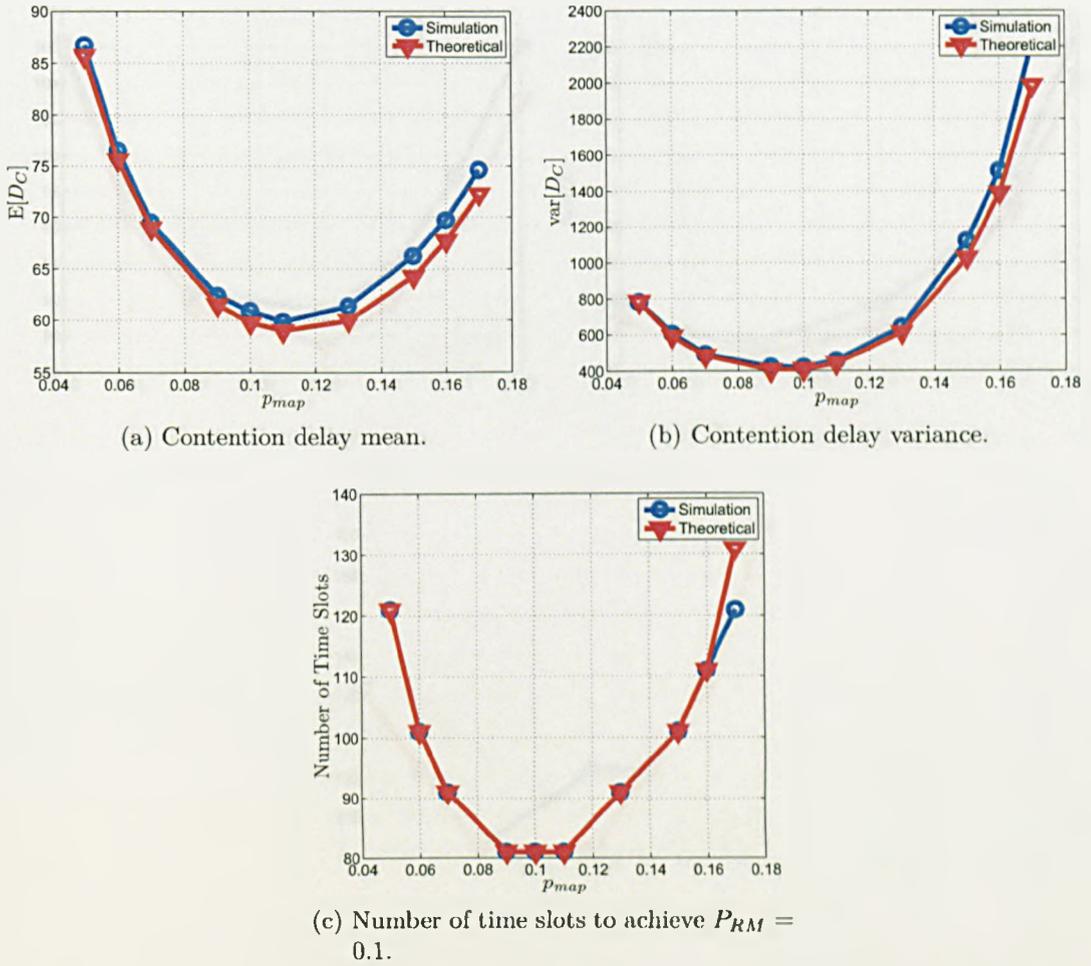


Figure 4.5.3: Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 1 \times 10^{-3}$ and $\gamma = 172$ ($P_{fa} = 10^{-5}$).

4.5. SIMULATION RESULTS AND DISCUSSION

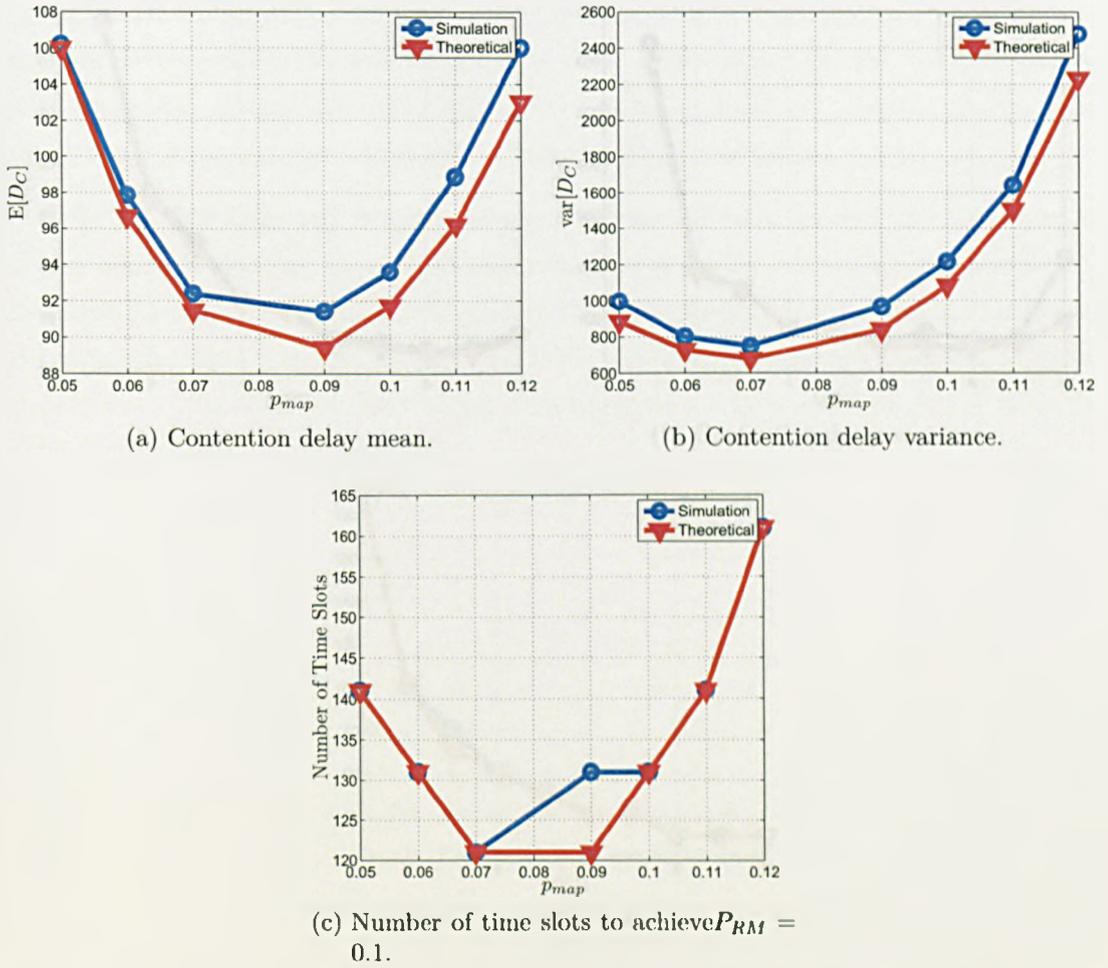


Figure 4.5.4: Contention delay statistics plotted against the medium access probability (p_{map}) with $\lambda = 5 \times 10^{-3}$ and $\gamma = 172$ ($P_{fa} = 10^{-5}$).

4.5. SIMULATION RESULTS AND DISCUSSION

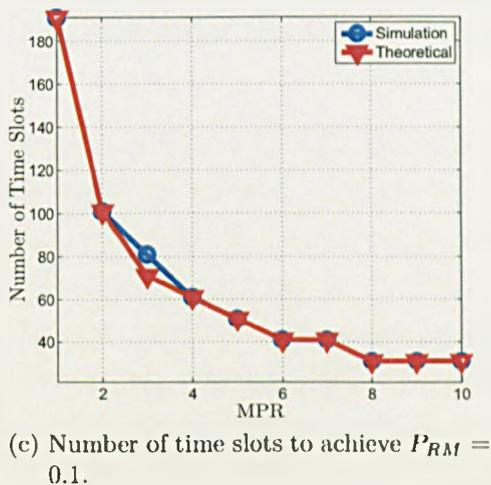
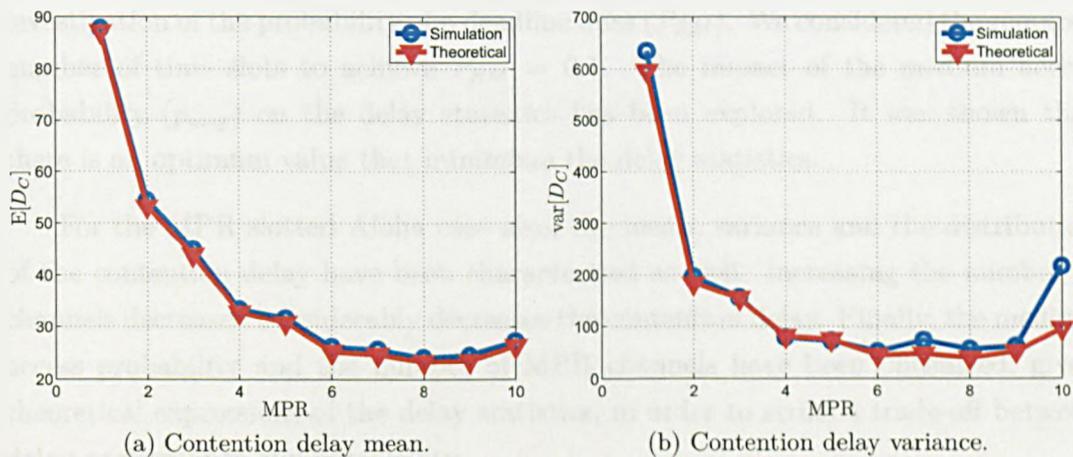


Figure 4.5.5: Contention delay statistics plotted against the number of MPR channels for $\lambda = 1 \times 10^{-2}$ and $\gamma = 150$ ($P_{fa} = 10^{-3}$).

4.6 Chapter Summary

We have defined and statistically characterized the contention delay (D_c) in a DDLE based WSN using the slotted Aloha MAC protocol with and without MPR. The stochastic geometry framework has been used here as well to aid in performance analysis. We have found the theoretical expressions for the mean and variance of the contention delay. The latter were in turn used in the moment matching approximation of the contention delay distribution. Finding the distribution enabled investigation of the probability of a deadline miss (P_{RM}). We considered the required number of time slots to achieve $P_{RM} = 0.1$. The impact of the medium access probability (p_{map}) on the delay statistics has been explored. It was shown that there is an optimum value that minimizes the delay statistics.

For the MPR slotted Aloha case also, the mean, variance and the distribution of the contention delay have been characterized as well. Increasing the number of channels decreased considerably decreases the contention delay. Finally, the medium access probability and the number of MPR channels have been optimized, given theoretical expressions of the delay statistics, in order to strike a trade-off between delay performance and complexity.

Chapter 5

Overview, Conclusions, and Future Work

In this thesis, we have considered the problem of distributed detection, localization and estimation (DDLE) of a stationary target using a WSN subjected to a time-critical operation, restricted power and BW resources, and a Rayleigh fading channel.

To this end, we proposed a modular DDLE algorithm that relies on two main stages. In the first stage, resource efficient distributed detection is performed to decide whether a target is present or not. The proposed distributed detection comprises from energy local detection carried out in the SNs and counting rule fusion performed at the FC. This structure does not require explicit knowledge about the target location or other parameters, which simplifies the distributed detection. Furthermore, it enables censoring the SNs to preserve resources and using a shared channel for communication. The effect of fading on distributed detection is handled by using the novel distributed diversity combining technique. We propose two distributed techniques: the distributed Maximum Ratio Combining (dMRC) and the distributed Equal Gain Combining (dEGC). The detection performance improves with increasing λ . There is an optimal γ value that minimizes the probability of error.

In the second stage, target localization and estimation is implemented using the novel segmented distributed localization and estimation (SDLE). The localization and estimation problem is decoupled into a group of local estimation problems solved at the SNs and a global estimation problem, which delivers the final estimate, solved

5.1. CONCLUSIONS

at the FC. We proposed two algorithms for solving the local problem: a nonlinear least squares (NLS) algorithm using the variable projection (VP) method and the grid search (GS) method. We proposed four algorithms to solve the global problem: NLS, GS, , hyperspherical intersection method (HSI) and the robust hyperspherical intersection (RHSI) method. Thus, the SDLE can be solved through various local and global algorithm combinations. We tried out five combinations: NLS², NLS-HSI, NLS-RHSI, GS², and GS-NLS. It turned out that the last algorithm combination delivers the best localization and estimation performance. In fact, the target can be localized with less than one meter error. Intuitively, increasing λ improves the MSE. However the MSE is insensitive to variations in the γ values used.

The local estimate transmission is done over a shared channel to save resources. Hence we propose using the slotted Aloha protocol. However, due to the relatively high delay in the latter protocol we adopt using multi-packet reception (MPR) with the slotted Aloha. We have found the theoretical expressions for the mean and variance of the contention delay. The latter was in turn used in the moment matching approximation of the contention delay distribution. Given theoretical expressions of the delay statistics enables optimizing the medium access probability and the number of MPR channels in order to strike a trade-off between delay performance and complexity.

5.1 Conclusions

It has been shown in practical surveillance WSN systems that the signal processing algorithm complexity can be compensated by the sheer number of SNs involved in the overall DDLE task. In other words, a large number of SNs with simple local processing can perform the same DDLE task (more or less) as a relatively small number of SN equipped with complex local signal processing. In this thesis, we have managed to start with a complex local signal processing algorithm (GLRT) and break it up into a group of simpler problems solved locally and globally.

For example, in the case of distributed detection, it has been shown that the distributed diversity combining techniques (dMRC and dEGC) performed better than centralized diversity combining even with high diversity order under low and medium SNRs. Because in the former, each SNs increases the diversity gain in contrast to the later that uses dedicated MACs to increase diversity. In fact, performance levels close to the ideal case can be reached through the use distributed diversity combining because the channel effects are eliminated and the network can be then turned into a distributed beamformer.

5.1. CONCLUSIONS

The performance of the distributed detection can be controlled through the system parameters λ and γ (deployment density and local detection threshold). In other words, the distributed detection system can be designed by first choosing P_D and P_{FA} values. From P_{FA} the global detection threshold (Υ) is readily found. Then, the λ and γ parameters are chosen to satisfy the P_D requirement. Furthermore, low P_e can be achieved while on the mean time turning off a significant number of SNs to save energy. This can be done by appropriately choosing the local detection threshold.

As for distributed estimation, breaking up the main MLE problem resulted in a group of simpler local problems that can be solved simply by using grid search techniques that reduce the required hardware to a simple correlator, while in the same time saving BW and energy by transmitting the intermediate estimates instead of the raw measurement. Results showed that a localization error of less than one meter can be achieved in a sensing field of $100m \times 100m$ area. Furthermore, such paradigm enables self censoring of the nodes that believes that their data are not good enough, which extends the system's lifetime. Interestingly, when decreasing the number of involved SNs in the estimation problem, the performance is marginally affected. For example, 14% reduction in the number of detecting SNs can be achieved by increasing γ , whereas the loss of localization performance is negligible in the case of the NLS² algorithm.

Furthermore, the symmetry between the local and global processing enables the deployment of a homogenous network. The proposed algorithm local and global parts both consist of simple detection operation (ED and CR respectively) followed by the estimation operation, which can be solved via iterative or grid search methods. Thus, virtually the same algorithms are used to solve the local and global estimation problem. Hence, any SN can fulfill the duties of the FC.

Generally speaking, the performance of the proposed DDLE algorithm mainly depends on the system parameters (λ , γ , p_{map} , and L), which are the SN deployment density, the local detection threshold, medium access probability, and number of MPR channels respectively. The joint optimization of all the previous parameters is a challenging task. On one hand, the performance measures are diverse, i.e., error probability (P_e) in the case of distributed detection, localization and estimation means square errors (MSEs) in the case of distributed estimation, contention delay in case of slotted Aloha, and number of channels for the MPR. On the other hand, the joint optimization problem is high dimensional and irregular. For example, increasing λ improves both distributed detection and estimation. However it increases the contention delay of the system. Therefore, we have suggested taking a

5.2. FUTURE WORK

sequential approach in which each parameter is optimized individually while fixing the others. For example, given a particular P_e and MSE, λ should be first chosen to satisfy both. Then, γ should be increased (which reduces the number of involved SNs and consequently preserves power and BW) to the maximum value that still maintains P_e . Then, L and p_{map} are chosen to reduce the contention delay while at the same time keeping the SN hardware within reasonable complexity.

In conclusion, the proposed algorithm enjoys the following features that are suitable for WSNs operation:

1. Low bandwidth requirement due to sending the local decision and intermediate estimates instead of the raw measurement.
2. Low power requirement and consequently extended network lifetime, since a portion of the SNs are not involved in the DDLE process.
3. Low computational requirement, e.g., due to using ED for local detection and the GS processing, which is simply a correlator.

However, the above advantages come with the price of losing optimality and increased estimation variance. Nonetheless, the performance loss can be compensated by proper choice of λ and γ .

5.2 Future Work

The future work can be divided into two main categories, short term and long term. In the short term part, our work can be directly extended to new cases. In the long term part, the suggested algorithms and techniques can be used to solve different problems.

5.2.1 Short term Work

5.2.1.1 Distributed Detection

We can consider using more sophisticated local detectors instead of the ED. On the other hand, channel estimation error could be incorporated into the dMRC and dEGC algorithms to investigate its effect. Moreover, statistical channel information could be used instead of the actual realization. More optimistic channel fading models could be used, such as Rician fading, etc.

5.2. FUTURE WORK

5.2.1.2 Distributed Estimation

In the SDLE framework, different local or global estimation algorithms could be investigated. Similarly, in the case of global estimation, an optimal estimator can be derived given that the local intermediate parameter statistical nature is identified. On a different note, the effect of quantization of the intermediate parameters can be investigated. And also consider the effect of reception errors of those parameters.

5.2.1.3 MAC Protocol

The CSMA or SMAC protocols could be used instead of slotted Aloha, and the delay statistic could be analyzed using the same proposed framework. A lower bound on the contention delay (in the case of the MPR) could also be investigated.

5.2.2 Long term Work

For long term work, different models can be considered in lieu of the diffusive source model, such as the wave field source, which can represent acoustic or electromagnetic targets. It is expected though, that the SDLE analysis would give more insight since the wave field source model is more tractable. Furthermore, a moving target tracking can be investigated using the DDLE framework, particularly, through the concept of the intermediate parameters used in Chapter 3. In this case, the tracking problem can be handled locally and globally via a framework similar to the SDLE. The DDLE can be used in the multiple target scenario, in which spatial clustering can be used with several instances of the DDLE algorithm. WSN specific simulators (such as TOSSIM [135]) can also be used to investigate the effect of different communication aspects.

Appendix A

Proofs in Chapter 2

A.1 Proof of Theorem 2.1

Under hypothesis \mathcal{H}_0 , the statistics $\Xi(\mathbf{X}_i) \forall \mathbf{X}_i \in \Phi$ do not depend on the target, in fact they are independently and identically distributed since they are only functions of noise and γ . Thus, the probability of retaining that point in Φ_d^0 , i.e., the thinning probability of a point at \mathbf{X}_i , is

$$\mathbb{P}(\mathbf{X}_i \in \Phi_d) = \mathbb{P}(\Xi(\mathbf{X}_i) \geq \gamma; \mathcal{H}_0) \quad (\text{A.1.1})$$

which is constant and equals the local probability of false alarm, say P_{fa} . Hence, the detecting SNs then form the point process $\Phi_d^0 = \{\mathbf{X}_i \in \Phi : \Xi(\mathbf{X}_i) \geq \gamma\}$ that is a p -thinning of Φ . Such operation produces a homogenous PPP. The mean is simply $\lambda P_{fa} |\mathcal{F}|$. As a result, Ω_0 's distribution follows a Poisson distribution given in (2.3.3).

A.2 Proof of Theorem 2.2

Since Φ is stationary and \mathcal{F} is infinitely large, we construct the shifted MPPP $\Phi' = \{\mathbf{X}_i, \Xi'(\mathbf{X}_i), H(\mathbf{X}_i), \Psi(\mathbf{X}_i) : \mathbf{X}_i \in \Phi - \mathbf{X}_t\}$. The new MPPP is centered at \mathbf{X}_t , hence eliminating the dependence on the target's position, which will significantly simplify the proof. We first prove that Φ_d^1 is a PPP by noting that Φ_d^1 is a *dependently thinned* point process of Φ' , i.e., $\Phi_d^1 = \{\mathbf{X}_i \in \Phi' : \Xi(\mathbf{X}_i, A) \geq \gamma\}$, and hence it is a PPP. Note that the notation $\Xi(\mathbf{X}_i, A)$ is used here instead of $\Xi(\mathbf{X}_i)$ to make clear

the dependance on both \mathbf{X}_i and A . To show that Φ_d^1 is inhomogeneous, we examine its intensity by first finding its mean:

$$\begin{aligned}\Omega_1 &= \sum_{\mathbf{x}_i \in \Phi_d^1} I(\mathbf{X}_i) \\ &= \sum_{\mathbf{x}_i \in \Phi} \mathbf{1}(\Xi(\mathbf{X}_i, A) \geq \gamma)\end{aligned}\tag{A.2.1}$$

where $\mathbf{1}(\cdot)$ is the indicator function.

Let us condition Ω_1 on a specific value of the target's amplitude say a , then the conditioned mean number of detecting SNs, say $\Delta(a)$, is

$$\Delta(a) = \mathbb{E}_1[\Omega|a] = \mathbb{E}_1\left[\sum_{\mathbf{x}_i \in \Phi'} \mathbf{1}(\Xi(\mathbf{X}_i, a) \geq \gamma)\right]\tag{A.2.2}$$

where the expectation is w.r.t to Φ' . For an arbitrary mark cumulative distribution $P(\xi)$, Campbell's theorem yields the following

$$\begin{aligned}\Delta(a) &= \lambda \int_{\mathcal{F}} \int_0^\infty \mathbf{1}(\Xi(\mathbf{x}, a) > \gamma; \mathcal{H}_1) dP(\xi) d\mathbf{x} \\ &= \lambda \int_{\mathcal{F}} \mathbb{P}(\Xi(\mathbf{x}, a) \geq \gamma; \mathcal{H}_1) d\mathbf{x} \\ &= \lambda \int_{\mathcal{F}} P_d(\mathbf{x}, a) d\mathbf{x}.\end{aligned}\tag{A.2.3}$$

From the second line above, the intensity is $\lambda P_d(\mathbf{x}, a)$ which is dependent on \mathbf{x} . Hence we deduce that Φ_d^1 is a inhomogeneous point process. Hence, the conditional distribution of Ω_1 is $p(\omega|a) = \text{Poi}(\Delta(a))$. Consequently, Ω_1 's distribution is $p_1(\theta) = \mathbb{E}_A[\text{Poi}(\Delta(A))]$, which is the Cox distribution. Finally, the mean of Ω_1 is attained by take the expectation of Δ w.r.t $p(a)$

$$\begin{aligned}\bar{\Omega}_1 &= \mathbb{E}_A[\Delta(A)] \\ &= \lambda \mathbb{E}\left[\int_{\mathcal{F}} P_d(\mathbf{x}, A) d\mathbf{x}\right].\end{aligned}\tag{A.2.4}$$

A.3 Proof of Lemma 2.1

Let us group the signal samples in the vector $\mathbf{F}_i = (F_i[0], \dots, F_i[2K-1])^T$ with realization $\mathbf{f}_i = (f_i[0], \dots, f_i[2K-1])^T$. Also let $\Xi_i = \Xi(\mathbf{F}_i)$ and $d_i = \|\mathbf{x}_i - \mathbf{x}_t\|$ for convenience. We begin, by identifying the following Markov chain $\mathbf{F}_i \rightarrow \Xi(\mathbf{F}_i)$ that relates the observed signal with the LLR of the local detector at SN located at \mathbf{X}_i . By using the data processing inequality for the J-divergence [88], we have

$$\begin{aligned} J(p_0(\xi_i), p_1(\xi_i)) &\leq J(p_0(\mathbf{f}_i), p_1(\mathbf{f}_i)) \\ &= \mathbb{E}_1 \left[\log \left(\frac{p_1(\mathbf{f}_i)}{p_0(\mathbf{f}_i)} \right) \right] - \mathbb{E}_0 \left[\log \left(\frac{p_1(\mathbf{f}_i)}{p_0(\mathbf{f}_i)} \right) \right]. \end{aligned} \quad (\text{A.3.1})$$

Secondly, we characterize the LLR

$$\begin{aligned} \log \left(\frac{p_1(\mathbf{f}_i)}{p_0(\mathbf{f}_i)} \right) &= \log \left(\frac{\mathbb{E}_A [p_1(\mathbf{f}_i|A)]}{p_0(\mathbf{f}_i)} \right) \\ &\leq \mathbb{E}_A \left[\log \left(\frac{p_1(\mathbf{f}_i|A)}{p_0(\mathbf{f}_i)} \right) \right]. \end{aligned} \quad (\text{A.3.2})$$

Then it follows that

$$\mathbb{E}_1 \left[\log \left(\frac{p_1(\mathbf{f}_i)}{p_0(\mathbf{f}_i)} \right) \right] \leq \mathbb{E}_1 \left[\mathbb{E}_A \left[\log \left(\frac{p_1(\mathbf{f}_i|A)}{p_0(\mathbf{f}_i)} \right) \right] \right] \quad (\text{A.3.3})$$

$$\mathbb{E}_0 \left[\log \left(\frac{p_1(\mathbf{f}_i)}{p_0(\mathbf{f}_i)} \right) \right] \leq \mathbb{E}_0 \left[\mathbb{E}_A \left[\log \left(\frac{p_1(\mathbf{f}_i|A)}{p_0(\mathbf{f}_i)} \right) \right] \right]. \quad (\text{A.3.4})$$

By adding the above inequalities we arrive at

$$J(p_0(\mathbf{f}_i), p_1(\mathbf{f}_i)) \leq J(p_0(\mathbf{f}_i|A), p_1(\mathbf{f}_i)). \quad (\text{A.3.5})$$

To evaluate the LHS of the above, we evaluate the LHS of both of the (A.3.3) and (A.3.4). Starting with the former, we need to evaluate the LLR

$$\begin{aligned}
\log(p_0(\mathbf{s}_i)) &= \log\left((\pi\sigma_s^2)^{-\frac{N}{2}} \exp\left(-\frac{1}{2\sigma_s^2} \sum_{n=0}^{N-1} f_i^2[n]\right)\right) \\
&= -\frac{N}{2} \log(\pi\sigma_s^2) - \frac{1}{2\sigma_s^2} \sum_{n=0}^{N-1} f_i^2[n]
\end{aligned} \tag{A.3.6}$$

and

$$\begin{aligned}
\log(p_1(\mathbf{s}_i)) &= -\frac{N}{2} \log(\pi\sigma_s^2) - \frac{1}{2\sigma_s^2} \sum_{n=0}^{N-1} \left(f_i[n] - \frac{A}{d_i}\right)^2 \\
&= -\frac{N}{2} \log(\pi\sigma_s^2) - \frac{NA^2}{2d_i^2\sigma_s^2} + \frac{1}{\sigma_s^2} \sum_{n=0}^{N-1} \frac{Af_i[n]}{d_i} - \frac{1}{2\sigma_s^2} \sum_{n=0}^{N-1} f_i^2[n].
\end{aligned} \tag{A.3.7}$$

Then we have

$$\begin{aligned}
\mathbb{E}_1 \left[\mathbb{E}_A \left[\log \left(\frac{p_1(\mathbf{f}_i|A)}{p_0(\mathbf{f}_i)} \right) \right] \right] &= \mathbb{E}_A \left[\mathbb{E}_1 \left[\log \left(\frac{p_1(\mathbf{f}_i|A)}{p_0(\mathbf{f}_i)} \right) \right] \right] \\
&= \mathbb{E}_A \left[\mathbb{E}_1 \left[-\frac{1}{2\sigma_s^2} \left(\frac{NA^2}{d_i^2} - 2 \sum_{n=0}^{N-1} \frac{A}{d_i} \left(\frac{A}{d_i} + w_i[n] \right) \right) \right] \right] \\
&= \frac{N\mathbb{E}[A^2]}{2\sigma_s^2 d_i^2}.
\end{aligned} \tag{A.3.8}$$

Similarly

$$\mathbb{E}_0 \left[\mathbb{E}_A \left[\log \left(\frac{p_1(\mathbf{f}_i|A)}{p_0(\mathbf{f}_i)} \right) \right] \right] = -\frac{N\mathbb{E}[A^2]}{2\sigma_s^2 d_i^2}. \tag{A.3.9}$$

Consequently, by subtracting (A.3.9) from (A.3.8) we finally arrive at

$$J(p_0(\mathbf{f}_i|A), p_1(\mathbf{f}_i)) = \frac{N\mathbb{E}[A^2]}{\sigma_s^2 d_i^2}. \tag{A.3.10}$$

A.4 Proof of Proposition 2.1

Given a particular value for the target's amplitude, i.e., $A = a$, the approximate mean is

$$\begin{aligned}
\Delta &\approx \lambda \int_{\mathcal{B}} 1 \times r' dr' + \lambda \int_{\bar{\mathcal{B}}} P_{f_a} r' dr' \\
&= \lambda \int_0^r r' dr' + \lambda P_{f_a} \int_r^{f(\bar{\mathcal{B}})} r' dr'
\end{aligned} \tag{A.4.1}$$

where $f(\bar{\mathcal{B}})$ the function defining the outer boundary of $\bar{\mathcal{B}}$. Using the result in lemma 2.1, for arbitrary ϵ , we set $r = \frac{a}{\sigma_s} \sqrt{\frac{N}{\epsilon}}$ and then substituting into (A.4.1) we get

$$\begin{aligned}
\Delta &\approx \frac{\lambda N a^2}{2\epsilon\sigma_s^2} + \frac{\lambda P_{f_a}}{2} \left(f^2(\bar{\mathcal{B}}) - \frac{N a^2}{\epsilon\sigma_s^2} \right) \\
&= \frac{\lambda N a^2 (1 - P_{f_a})}{2\epsilon\sigma_s^2} + \frac{\lambda P_{f_a} f^2(\bar{\mathcal{B}})}{2}.
\end{aligned} \tag{A.4.2}$$

Hence, Δ is approximately a quadratic function w.r.t a . Consequently, there exist some constants c_0 , c_1 , and c_2 such that

$$\Delta \approx g(a) = c_0(a - c_1)^2 + c_2. \tag{A.4.3}$$

The above is simply a transformation of RV A to Δ . Hence the its distribution follows as in (2.3.15).

A.5 Proof of Proposition 2.2

Using the same translation argument in proof A.2, the target location dependence is eliminated. Now note that $f_1(\mathbf{x})$ is always non-negative given the condition $P_d(\mathbf{x}) \geq P_{f_a}$ is satisfied for all \mathbf{x} . Then, Λ_{OFR}^1 is a summation of a non-negative function over Φ_d , which has intensity of $\lambda P_d(\mathbf{x})$ as stated in corollary (2.2). Hence, Campbell's theorem can be readily used to find the mean as follows

$$\mathbb{E}[\Lambda_{\text{OFR}}^1] = \mathbb{E}\left[\sum_{\mathbf{X}_i \in \Phi_d} f(\mathbf{X}_i)\right] = \lambda \int_{\mathcal{F}} f(\mathbf{x}) P_d(\mathbf{x}) d\mathbf{x}. \tag{A.5.1}$$

Similarly, the variance is given by

$$\text{var} \left[\Lambda_{\text{OFR}}^1 \right] = \text{var} \left[\sum_{\mathbf{x}_i \in \Phi_d} f(\mathbf{X}_i) \right] = \lambda \int_{\mathcal{F}} f^2(\mathbf{x}) P_d(\mathbf{x}) d\mathbf{x}. \quad (\text{A.5.2})$$

For Λ_{OFR}^0 though, we apply the same procedure but this time $\overline{\Phi}_d$ has intensity $\lambda(1 - P_d(\mathbf{x}))$. As for the \mathcal{H}_0 case, the proof is similar. However the intensities of Φ_d and $\overline{\Phi}_d$ are λP_{f_a} and $\lambda(1 - P_{f_a})$ respectively.

A.6 Proof of Lemma 2.2

Starting with the LLR and using Jensen's inequality

$$\begin{aligned} \log \left(\frac{p_1(y)}{p_0(y)} \right) &= \log \left(\frac{\mathbb{E}_1 \left[\exp \left(\frac{-1}{2\sigma_c^2} (Y - \Omega)^2 \right) \right]}{\mathbb{E}_0 \left[\exp \left(\frac{-1}{2\sigma_c^2} (Y - \Omega)^2 \right) \right]} \right) \\ &\leq \frac{1}{2\sigma_c^2} \mathbb{E}_0 \left[(Y - \Omega)^2 \right] - \frac{1}{2\sigma_c^2} \mathbb{E}_1 \left[(Y - \Omega)^2 \right] \\ &= \frac{y}{2\sigma_c^2} (\mathbb{E}_1[\Omega] - \mathbb{E}_0[\Omega]) - \frac{1}{2\sigma_c^2} (\mathbb{E}_1[\Omega^2] - \mathbb{E}_0[\Omega^2]). \end{aligned} \quad (\text{A.6.1})$$

Now taking the difference of the expectation w.r.t \mathcal{H}_1 and \mathcal{H}_0 we get the inequality

$$\begin{aligned} J(p_0(y), p_0(y)) &\leq \frac{1}{2\sigma_c^2} (\mathbb{E}_1[Y] - \mathbb{E}_0[Y]) (\mathbb{E}_1[\Omega] - \mathbb{E}_0[\Omega]) \\ &= \frac{1}{2\sigma_c^2} (\mathbb{E}_1[\Omega] - \mathbb{E}_0[\Omega])^2 \end{aligned} \quad (\text{A.6.2})$$

where the last equality results in having $\mathbb{E}_j[Y] = \mathbb{E}_j[\Omega]$, which follows from (2.6.11).

Appendix B

Derivation and Proofs in Chapter 3

B.1 Analytical Expressions for Derivatives in the Diffusion Case

Recall that

$$h[n, \theta] = \frac{\mu}{4\pi\kappa\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}} \operatorname{erfc}\left(\frac{\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}}{2\sqrt{\kappa T_s(n - n_\tau)}}\right). \quad (\text{B.1.1})$$

Then the derivatives w.r.t $\theta = (\mu, x_t, y_t, n_\tau, \kappa)$ are

$$\frac{\partial h[n, \theta]}{\partial \mu} = \frac{1}{4\pi\kappa\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}} \operatorname{erfc}\left(\frac{\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}}{2\sqrt{\kappa T_s(n - n_\tau)}}\right). \quad (\text{B.1.2})$$

$$\begin{aligned} \frac{\partial h[n, \theta]}{\partial x_t} &= \frac{\mu}{4\pi\kappa[(x_i - x_t)^2 + (y_i - y_t)^2]^{\frac{3}{2}}} \operatorname{erfc}\left(\frac{\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}}{2\sqrt{\kappa T_s(n - n_\tau)}}\right) \\ &\quad - \frac{\mu(x_i - x_t)}{4(\pi\kappa)^{\frac{3}{2}}((x_i - x_t)^2 + (y_i - y_t)^2)\sqrt{T_s(n - n_\tau)}} \exp\left(-\frac{(x_i - x_t)^2 + (y_i - y_t)^2}{4\kappa T_s(n - n_\tau)}\right). \end{aligned} \quad (\text{B.1.3})$$

$$\frac{\partial h[n, \theta]}{\partial n_\tau} = \frac{\mu}{8T_s^2[\pi\kappa(n - n_\tau)]^{\frac{3}{2}}} \exp\left(-\frac{(x_i - x_t)^2 + (y_i - y_t)^2}{4\kappa T_s(n - n_\tau)}\right). \quad (\text{B.1.4})$$

$$\begin{aligned} \frac{\partial h[n, \theta]}{\partial \kappa} &= -\frac{\mu}{4\pi\kappa\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}} \operatorname{erfc}\left(\frac{\sqrt{(x_i - x_t)^2 + (y_i - y_t)^2}}{2\sqrt{\kappa T_s(n - n_\tau)}}\right) \\ &\quad - \frac{\mu}{8\pi^{\frac{3}{2}}\kappa^{\frac{3}{2}}\sqrt{T_s(n - n_\tau)}} \exp\left(-\frac{(x_i - x_t)^2 + (y_i - y_t)^2}{4\kappa T_s(n - n_\tau)}\right). \end{aligned} \quad (\text{B.1.5})$$

The derivatives with respect to y_t and z_t are similar to (B.1.3).

B.2 Derivation of Fisher Information Matrix for an arbitrary Sensor Node

Recall that the individual LLF

$$\begin{aligned} l(\mathbf{x}_i, \theta) &= \frac{1}{2\sigma_i^2} \sum_{n=0}^{N-1} (f_i[n] - \mu h[n, \mathbf{x}_i, \psi])^2 \\ &= \frac{1}{2} \sum_{n=0}^{N-1} q_i^2[n, \theta]. \end{aligned} \quad (\text{B.2.1})$$

The individual FIM follows as

$$\mathcal{I}(\mathbf{x}_i, \theta) = -\mathbb{E}_w \left[\frac{\partial^2}{\partial \theta^2} \sum_{n=0}^{N-1} q_i^2[n, \theta] \right]. \quad (\text{B.2.2})$$

Since the individual LLF, $l(\mathbf{x}_i, \theta)$, has a sum of squares form, the above takes the following form

$$\mathcal{I}(\mathbf{x}_i, \boldsymbol{\theta}) = -\mathbb{E}_w \left[\mathbf{J}_i^T(\boldsymbol{\theta}) \mathbf{J}_i(\boldsymbol{\theta}) \right] - \mathbb{E}_w \left[\sum_{n=0}^{N-1} q_i[n, \boldsymbol{\theta}] \frac{\partial^2}{\partial \boldsymbol{\theta}^2} q_i[n, \boldsymbol{\theta}] \right] \quad (\text{B.2.3})$$

where $\mathbf{J}_i(\boldsymbol{\theta})$ is the Jacobian matrix defined as

$$\begin{aligned} \mathbf{J}_i(\boldsymbol{\theta}) &= \frac{\partial}{\partial \boldsymbol{\theta}} \mathbf{q}_i(\boldsymbol{\theta}) \\ &= \frac{1}{\sigma_s^2} (\mathbf{f}_i - \mu \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}))^T \left(\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \quad \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \right) \\ &= \frac{1}{\sigma_s^2} \mathbf{w}_i^T \left(\mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \quad \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \right) \end{aligned} \quad (\text{B.2.4})$$

where $\mathbf{q}_i(\boldsymbol{\theta}) = (q_i[0, \boldsymbol{\theta}], \dots, q_i[N-1, \boldsymbol{\theta}])^T$, $\mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi})$ is the derivative with respect to $\boldsymbol{\psi}$ as defined in (3.4.8), and the last equality above follows when $\boldsymbol{\theta}$ takes its true value. In this case the second term of (B.2.3) vanishes. Hence (B.2.3) becomes

$$\begin{aligned} \mathcal{I}(\mathbf{x}_i, \boldsymbol{\theta}) &= -\frac{1}{\sigma_s^4} \mathbb{E}_w \left[\begin{array}{cc} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{w}_i \mathbf{w}_i^T \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) & \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{w}_i \mathbf{w}_i^T \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \\ \mathbf{D}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{w}_i \mathbf{w}_i^T \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) & \mathbf{D}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{w}_i \mathbf{w}_i^T \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \end{array} \right] \\ &= -\frac{1}{\sigma_s^4} \left(\begin{array}{cc} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbb{E}_w \left[\mathbf{w}_i \mathbf{w}_i^T \right] \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) & \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbb{E}_w \left[\mathbf{w}_i \mathbf{w}_i^T \right] \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \\ \mathbf{D}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbb{E}_w \left[\mathbf{w}_i \mathbf{w}_i^T \right] \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) & \mathbf{D}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbb{E}_w \left[\mathbf{w}_i \mathbf{w}_i^T \right] \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \end{array} \right) \\ &= -\frac{1}{\sigma_s^2} \left(\begin{array}{cc} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) & \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \\ \mathbf{D}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) & \mathbf{D}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}) \end{array} \right). \end{aligned} \quad (\text{B.2.5})$$

B.3 Proof of Theorem 3.1

Starting from (3.4.1) and (3.4.2) we have

$$\text{CRLB}(\boldsymbol{\theta}) = - \left(\mathbb{E} \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^2} \right] \right)^{-1}. \quad (\text{B.3.1})$$

Recalling that $\mathbb{E}_w [q_i[n]] = 0$ in equation (B.2.3) gives

$$\mathbb{E} \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} \right] = \mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right) \right] \quad (\text{B.3.2})$$

which can be written as the following block matrix

$$\mathbb{E} \left[\frac{\partial^2 l(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}^T \partial \boldsymbol{\theta}} \right] = \mathbb{E} \begin{bmatrix} \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T & \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right) & \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T & \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \psi} \right) \\ \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T & \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right) & \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \psi} \right)^T & \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \psi} \right) \end{bmatrix}. \quad (\text{B.3.3})$$

We proceed by finding the expectations of the matrix elements one by one. For the upper left element we have

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta})}{\partial \mu} &= \frac{1}{\sigma_s^2} \sum_{\mathbf{x}_i \in \Phi} (\mathbf{f}_i - \mu \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}))^T \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \\ &= \frac{1}{\sigma_s^2} \sum_{\mathbf{x}_i \in \Phi} \mathbf{w}_i^T \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \end{aligned} \quad (\text{B.3.4})$$

where the last equality holds at the true value of $\boldsymbol{\theta}$. Then, the expectation is

$$\begin{aligned} \mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right) \right] &= \frac{1}{\sigma_s^4} \mathbb{E} \left[\sum_{\mathbf{x}_i \in \Phi} \sum_{\mathbf{x}_j \in \Phi} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{w}_i \mathbf{w}_j^T \mathbf{h}(\mathbf{x}_j, \boldsymbol{\psi}) \right] \\ &= \frac{1}{\sigma_s^4} \mathbb{E}_{\Phi} \left[\sum_{\mathbf{x}_i \in \Phi} \sum_{\mathbf{x}_j \in \Phi} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbb{E}_w [\mathbf{w}_i \mathbf{w}_j^T] \mathbf{h}(\mathbf{x}_j, \boldsymbol{\psi}) \right] \end{aligned} \quad (\text{B.3.5})$$

where the last line comes from the fact that the measurement noise and the point process are independent. Also, using the spatio-temporal independence property of the measurement noise i.e.,

$$\mathbb{E} [\mathbf{w}_i \mathbf{w}_j^T] = \begin{cases} \sigma_s^2 \mathbf{I}, & i = j \\ 0, & i \neq j \end{cases} \quad (\text{B.3.6})$$

we get

$$\mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right) \right] = \frac{1}{\sigma_s^2} \mathbb{E}_{\Phi} \left[\sum_{\mathbf{x}_i \in \Phi} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \right]. \quad (\text{B.3.7})$$

Now since $\mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \geq 0$, Campbell's theorem can be used to finally arrive at

$$\mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right) \right] = \frac{\lambda}{\sigma_s^2} \int_{\mathbb{R}^2} \|\mathbf{h}(\mathbf{x}, \boldsymbol{\psi})\|^2 d\mathbf{x}. \quad (\text{B.3.8})$$

If censoring is used, the summation is carried out over Φ_d and the above becomes

$$\mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\psi}} \right) \right] = \frac{\lambda}{\sigma_s^2} \int_{\mathbb{R}^2} \|\mathbf{h}(\mathbf{x}, \boldsymbol{\psi})\|^2 P_d(\mathbf{x}, \gamma) d\mathbf{x}. \quad (\text{B.3.9})$$

Moving on to the second element in (B.3.3), we proceed as before to find the partial derivative

$$\begin{aligned} \frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\psi}} &= \frac{\mu}{\sigma_s^2} \sum_{\mathbf{x}_i \in \Phi} (\mathbf{f}_i - \mu \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}))^T \frac{\partial}{\partial \boldsymbol{\psi}} \mathbf{h}(\mathbf{x}_i, \boldsymbol{\psi}) \\ &= \frac{\mu}{\sigma_s^2} \sum_{\mathbf{x}_i \in \Phi} \mathbf{w}_i^T \mathbf{D}(\mathbf{x}_i, \boldsymbol{\psi}). \end{aligned} \quad (\text{B.3.10})$$

The expected value of the block element follows similarly as before

$$\begin{aligned} \mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \mu} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\psi}} \right) \right] &= \frac{\mu}{\sigma_s^4} \mathbb{E} \left[\sum_{\mathbf{x}_i \in \Phi} \sum_{\mathbf{x}_j \in \Phi} \mathbf{h}^T(\mathbf{x}_i, \boldsymbol{\psi}) \mathbf{w}_i \mathbf{w}_j^T \mathbf{D}(\mathbf{x}_j, \boldsymbol{\psi}) \right] \\ &= \frac{\lambda \mu}{\sigma_s^2} \int_{\mathbb{R}^2} \mathbf{h}^T(\mathbf{x}, \boldsymbol{\psi}) \mathbf{D}(\mathbf{x}, \boldsymbol{\psi}) P_d(\mathbf{x}, \gamma) d\mathbf{x} \end{aligned} \quad (\text{B.3.11})$$

given that $\mathbf{h}^T(\mathbf{x}, \boldsymbol{\varphi}) \mathbf{D}(\mathbf{x}, \boldsymbol{\varphi}) \geq 0$. And finally, the expectation of the last element in (B.3.3) is

$$\mathbb{E} \left[\left(\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\psi}} \right)^T \left(\frac{\partial l(\boldsymbol{\theta})}{\partial \boldsymbol{\psi}} \right) \right] = \frac{\lambda \mu}{\sigma_s^2} \int_{\mathbb{R}^2} \|\mathbf{D}(\mathbf{x}, \boldsymbol{\psi})\|^2 P_d(\mathbf{x}, \gamma) d\mathbf{x}. \quad (\text{B.3.12})$$

Aggregating (B.3.9), (B.3.11), and (B.3.12) gives us the FIM

$$\mathcal{I}(\boldsymbol{\theta}) = \begin{pmatrix} \tilde{\mathbf{h}}(\boldsymbol{\psi}, \gamma) & \tilde{\mathbf{G}}(\boldsymbol{\psi}, \gamma) \\ \tilde{\mathbf{G}}^T(\boldsymbol{\psi}, \gamma) & \tilde{\mathbf{D}}(\boldsymbol{\psi}, \gamma) \end{pmatrix}. \quad (\text{B.3.13})$$

Finally, using the matrix inversion lemma we arrive at the CRLB expressions in (3.4.13) and (3.4.14).

B.4 Proof of Lemma 3.1

In order to prove convexity of (3.5.23), it suffice to show that the Hessian matrix of (3.5.23) is positive semidefinite. The cost function of problem (3.5.23) is

$$C = \|\mathbf{p}(\boldsymbol{\varphi})\|^2 \quad (\text{B.4.1})$$

and recall that

$$\mathbf{p}(\boldsymbol{\varphi}) = \mathbf{P}_h^\perp(\boldsymbol{\varphi})\mathbf{f} \quad (\text{B.4.2})$$

where $\mathbf{P}_h^\perp(\boldsymbol{\varphi})$ is the projection matrix on the columns orthogonal to \mathbf{h} . The Hessian matrix then is

$$\frac{\partial^2 C}{\partial \boldsymbol{\varphi}^T \partial \boldsymbol{\varphi}} = \mathbf{J}^T(\boldsymbol{\varphi})\mathbf{J}(\boldsymbol{\varphi}) + \sum_{n=0}^{N-1} p[n, \boldsymbol{\varphi}] \frac{\partial^2}{\partial \boldsymbol{\varphi}^T \partial \boldsymbol{\varphi}} p[n, \boldsymbol{\varphi}] \quad (\text{B.4.3})$$

where recall that $\mathbf{J}(\boldsymbol{\varphi})$ is the Jacobian matrix defined as

$$\mathbf{J}(\boldsymbol{\varphi}) = \frac{\partial}{\partial \boldsymbol{\varphi}} \mathbf{p}(\boldsymbol{\varphi}). \quad (\text{B.4.4})$$

The first term in (B.4.3) is positive semidefinite according to the condition of the lemma. The lemma is proved if the second term in (B.4.3), however, is proved to vanish asymptotically and this is what is accomplished below:

$$\sum_{n=0}^{N-1} p[n, \boldsymbol{\varphi}] \frac{\partial^2}{\partial \boldsymbol{\varphi}^T \partial \boldsymbol{\varphi}} p[n, \boldsymbol{\varphi}] = \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} f[m]f[l] P_{nm}^\perp \frac{\partial^2}{\partial \boldsymbol{\varphi}^T \partial \boldsymbol{\varphi}} P_{nl}^\perp \quad (\text{B.4.5})$$

where P_{nm}^\perp is the element in n th row and m th column. Note that we can state that

$$\begin{aligned} f[m]f[l] &= w[m]w[l] + h[m, \boldsymbol{\varphi}]h[l, \boldsymbol{\varphi}] \\ &\approx \sigma_s^2 \delta[m-l] + h[m, \boldsymbol{\varphi}]h[l, \boldsymbol{\varphi}]. \end{aligned} \quad (\text{B.4.6})$$

Plugging (B.4.6) into (B.4.5) yields

$$\begin{aligned}
L &= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} \left(\sigma_s^2 \delta[m-l] + h[m, \varphi] h[l, \varphi] \right) P_{nm}^\perp \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp \\
&= \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} P_{nm}^\perp \sum_{l=0}^{N-1} \left(\sigma_s^2 \delta[m-l] + h[m, \varphi] h[l, \varphi] \right) \frac{\partial^2}{\partial \varphi^T \varphi^2} P_{nl}^\perp \\
&= \sum_{n=0}^{N-1} \left(\sum_{m=0}^{N-1} \sigma_s^2 P_{nm}^\perp \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp + \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} h[m, \varphi] h[l, \varphi] P_{nm}^\perp \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp \right) \\
&= \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} P_{nm}^\perp \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp + \underbrace{\sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \sum_{l=0}^{N-1} h[m, \varphi] h[l, \varphi] P_{nm}^\perp \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp}_{L'} \tag{B.4.7}
\end{aligned}$$

Taking the second term in the last equality, we show that

$$\begin{aligned}
L' &= \sum_{n=0}^{N-1} \sum_{l=0}^{N-1} h[l, \varphi] \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp \left(\sum_{m=0}^{N-1} P_{nm}^\perp h[m, \varphi] \right) \\
&= 0. \tag{B.4.8}
\end{aligned}$$

since

$$\sum_{m=0}^{N-1} P_{nm}^\perp h[m, \varphi] = 0 \tag{B.4.9}$$

by virtue of the complementary projection matrix. Now we have

$$\begin{aligned}
L &= \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} P_{nm}^\perp \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl}^\perp \\
&= \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} (1 - P_{nm}) \frac{\partial^2}{\partial \varphi^T \varphi} (1 - P_{nl}) \\
&= \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \frac{\partial^2}{\partial \varphi^2} (1 - P_{nl}) - P_{nm} \frac{\partial^2}{\partial \varphi^T \varphi} (1 - P_{nl}) \\
&= \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} P_{nm} \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl} - \frac{\partial^2}{\partial \varphi^T \varphi} P_{nl} \\
&= \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \left(\frac{\partial^2}{\partial \varphi^T \varphi} P_{nl} \right) (P_{nm} - 1). \tag{B.4.10}
\end{aligned}$$

Now using lemma 3.2 in [136] we can write the projection matrix as

$$\mathbf{P} = \frac{1}{KN} \mathbf{h}(\boldsymbol{\varphi}) \mathbf{h}^T(\boldsymbol{\varphi}) + o\left(\frac{1}{N}\right) \quad (\text{B.4.11})$$

where N here is the number of samples and

$$K = \int |h(t; \boldsymbol{\varphi})|^2 dt \quad (\text{B.4.12})$$

then (B.4.10) becomes

$$L \approx \sigma_s^2 \sum_{n=0}^{N-1} \sum_{m=0}^{N-1} \frac{1}{K^2 N^2} \left(h[n; \boldsymbol{\varphi}] h[l; \boldsymbol{\varphi}] - \frac{1}{KN} \right) \frac{\partial^2}{\partial \boldsymbol{\varphi}^T \partial \boldsymbol{\varphi}} h[n; \boldsymbol{\varphi}] h[l; \boldsymbol{\varphi}]. \quad (\text{B.4.13})$$

Finally, as $N \rightarrow \infty$ we have $L \rightarrow 0$. This concludes the proof.

References

- [1] C.-Y. Chong and S. P. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1247–1256, 2003.
- [2] L. F. Herrera-Quintero, F. Macia-Perez, H. Ramos-Morillo, and C. Lago-Gonzalez, "Wireless smart sensors networks, systems, trends and its impact in environmental monitoring," in *Communications, 2009. LATINCOM'09. IEEE Latin-American Conference on*. IEEE, 2009, pp. 1–6.
- [3] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *Communications Magazine, IEEE*, vol. 40, no. 8, pp. 102–114, 2002.
- [4] M. Vieira, J. Coelho, C.N., J. da Silva, D.C., and J. da Mata, "Survey on wireless sensor network devices," in *Emerging Technologies and Factory Automation, 2003. Proceedings. ETFA '03. IEEE Conference*, vol. 1, Sept. 2003, pp. 537–544.
- [5] N. Xu, "A survey of sensor network applications," *IEEE Communications Magazine*, vol. 40, 2002.
- [6] A. Mainwaring, D. Culler, J. Polastre, R. Szewczyk, and J. Anderson, "Wireless sensor networks for habitat monitoring," in *WSNA '02: Proceedings of the 1st ACM international workshop on Wireless sensor networks and applications*. New York, NY, USA: ACM, 2002, pp. 88–97.
- [7] G. Barrenetxea, F. Ingelrest, G. Schaefer, and M. Vetterli, "Wireless sensor networks for environmental monitoring: The sensorscope experience," in *Communications, 2008 IEEE International Zurich Seminar on*. IEEE, 2008, pp. 98–101.
- [8] S. eun Yoo, J. eon Kim, T. Kim, S. Ahn, J. Sung, and D. Kim, "A2s: Auto-

- mated agriculture system based on wsn,” in *Proc. IEEE International Symposium on Consumer Electronics ISCE 2007*, 2007, pp. 1–5.
- [9] G. Werner-Allen, J. Johnson, M. Ruiz, J. Lees, and M. Welsh, “Monitoring volcanic eruptions with a wireless sensor network,” in *Wireless Sensor Networks, 2005. Proceedings of the Second European Workshop on*. IEEE, 2005, pp. 108–120.
- [10] L. Krishnamurthy, R. Adler, P. Buonadonna, J. Chhabra, M. Flanigan, N. Kushalnagar, L. Nachman, and M. Yarvis, “Design and deployment of industrial sensor networks: experiences from a semiconductor plant and the north sea,” in *SenSys '05: Proceedings of the 3rd international conference on Embedded networked sensor systems*. New York, NY, USA: ACM Press, 2005, pp. 64–75.
- [11] F. Stajano, N. Hault, I. Wassell, P. Bennett, C. Middleton, and K. Soga, “Smart bridges, smart tunnels: Transforming wireless sensor networks from research prototypes into robust engineering infrastructure,” *Ad Hoc Networks*, vol. 8, no. 8, pp. 872 – 888, 2010.
- [12] T. Schmid, H. Dubois-Ferriere, and M. Vetterli, “Sensorscope: Experiences with a wireless building monitoring sensor network,” in *Workshop on real-world wireless sensor networks (REALWSN05)*, 2005, pp. 13–17.
- [13] M. V. Ramesh, “Real-time wireless sensor network for landslide detection,” in *SENSORCOMM '09: Proceedings of the 2009 Third International Conference on Sensor Technologies and Applications*. Washington, DC, USA: IEEE Computer Society, 2009, pp. 405–409.
- [14] K. Sha, W. Shi, and O. Watkins, “Using wireless sensor networks for fire rescue applications: Requirements and challenges,” in *Electro/information Technology, 2006 IEEE International Conference on*, May 2006, pp. 239–244.
- [15] Z. Sun, P. Wang, M. C. Vuran, M. A. Al-Rodhaan, A. M. Al-Dhelaan, and I. F. Akyildiz, “Bordersense: Border patrol through advanced wireless sensor networks,” *Ad Hoc Networks*, vol. 9, no. 3, pp. 468–477, 2011.
- [16] S. Kumar and D. Shepherd, “Sensit: Sensor information technology for the warfighter,” in *Proc. 4th Int. Conf. on Information Fusion*, 2001.
- [17] T. He, S. Krishnamurthy, J. A. Stankovic, T. Abdelzaher, L. Luo, R. Stoleru, T. Yan, L. Gu, J. Hui, and B. Krogh, “Energy-efficient surveillance system using wireless sensor networks,” in *MobiSys '04: Proceedings of the 2nd international conference on Mobile systems, applications, and services*. New York,

- NY, USA: ACM, 2004, pp. 270–283.
- [18] J. Hill and D. Culler, “Mica: a wireless platform for deeply embedded networks,” *Micro, IEEE*, vol. 22, no. 6, pp. 12–24, Nov 2002.
- [19] S. Oh, P. Chen, M. Manzo, and S. Sastry, “Instrumenting wireless sensor networks for real-time surveillance,” in *ICRA*, 2006, pp. 3128–3133.
- [20] P. Dutta, J. Hui, J. Jeong, S. Kim, C. Sharp, J. Taneja, G. Tolle, K. Whitehouse, and D. Culler, “Trio: enabling sustainable and scalable outdoor wireless sensor network deployments,” in *IPSN '06: Proceedings of the 5th international conference on Information processing in sensor networks*. New York, NY, USA: ACM, 2006, pp. 407–415.
- [21] A. Arora, R. Ramnath, E. Ertin, P. Sinha, S. Bapat, V. Naik, V. Kulathumani, H. Zhang, H. Cao, M. Sridharan *et al.*, “Exscal: Elements of an extreme scale wireless sensor network,” in *Embedded and Real-Time Computing Systems and Applications, 2005. Proceedings. 11th IEEE International Conference on*. IEEE, 2005, pp. 102–108.
- [22] A. Arora, P. Dutta, S. Bapat, V. Kulathumani, H. Zhang, V. Naik, V. Mittal, H. Cao, M. Demirbas, M. Gouda *et al.*, “A line in the sand: a wireless sensor network for target detection, classification, and tracking,” *Computer Networks*, vol. 46, no. 5, pp. 605–634, 2004.
- [23] G. Simon, M. Maróti, Ákos Lédeczi, G. Balogh, B. Kusy, A. Nádas, G. Pap, J. Sallai, and K. Frampton, “Sensor network-based countersniper system.” ACM Press, 2004, pp. 1–12.
- [24] Á. Lédeczi, P. Volgyesi, M. Maróti, G. Simon, G. Balogh, A. Nádas, B. Kusy, S. Dóra, and G. Pap, “Multiple simultaneous acoustic source localization in urban terrain,” in *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*. IEEE, 2005, pp. 491–496.
- [25] G. Balogh, A. Ledeczi, M. Maróti, and G. Simon, “Time of arrival data fusion for source localization,” in *Proceedings of The WICON Workshop on Information Fusion and Dissemination in Wireless Sensor Networks (SensorFusion 2005)*, 2005.
- [26] L. Gu, D. Jia, P. Vicaire, T. Yan, L. Luo, A. Tirumala, Q. Cao, T. He, J. A. Stankovic, T. Abdelzaher *et al.*, “Lightweight detection and classification for wireless sensor networks in realistic environments,” in *Proceedings of the 3rd international conference on Embedded networked sensor systems*. ACM, 2005, pp. 205–217.

- [27] I. Demirkol, C. Ersoy, and F. Alagoz, "Mac protocols for wireless sensor networks: a survey," *Communications Magazine, IEEE*, vol. 44, no. 4, pp. 115–121, 2006.
- [28] K. Akkaya and M. Younis, "A survey on routing protocols for wireless sensor networks," *Ad Hoc Networks*, vol. 3, no. 3, pp. 325–349, May 2005.
- [29] D. Li, K. D. Wong, Y. H. Hu, and A. M. Sayeed, "Detection, classification, and tracking of targets," *Signal Processing Magazine, IEEE*, vol. 19, no. 2, pp. 17–29, 2002.
- [30] R. R. Brooks, P. Ramanathan, and A. M. Sayeed, "Distributed target classification and tracking in sensor networks," *Proceedings of the IEEE*, vol. 91, no. 8, pp. 1163–1171, 2003.
- [31] P. Werr, "Blast rocks egypt's gas pipeline to israel, jordan,"
- [32] C. Rago, P. Willett, and Y. Bar-Shalom, "Censoring sensors: a low-communication-rate scheme for distributed detection," *IEEE Transactions On Aerospace And Electronic Systems*, vol. 32, no. 2, pp. 554–568, 1996.
- [33] J. Tsitsiklis, "Decentralized detection by a large number of sensors," *Mathematics of Control, Signals, and Systems (MCSS)*, vol. 1, pp. 167–182, 1988.
- [34] P. K. Varshney, *Distributed Detection and Data Fusion*, 1st ed. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 1996.
- [35] R. Viswanathan and P. Varshney, "Distributed detection with multiple sensors i. fundamentals," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 54–63, Jan. 1997.
- [36] R. Blum, S. Kassam, and H. Poor, "Distributed detection with multiple sensors i. advanced topics," *Proceedings of the IEEE*, vol. 85, no. 1, pp. 64–79, Jan. 1997.
- [37] R. Niu, P. K. Varshney, M. Moore, and D. Klammer, "Decision fusion in a wireless sensor network with a large number of sensors," in *In Fusion*, 2004, pp. 21–27.
- [38] R. Niu and P. K. Varshney, "Distributed detection and fusion in a large wireless sensor network of random size," *EURASIP Journal on Wireless Communication and Networking*, vol. 2005, no. 4, pp. 462–472, 2005.
- [39] 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, 2006.
- [40] R. Niu and P. Varshney, “Joint detection and localization in sensor networks based on local decisions,” in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, Nov. 2006, pp. 525 –529.
- [41] —, “Performance analysis of distributed detection in a random sensor field,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 339 –349, Jan. 2008.
- [42] D. Castanon and D. Teneketzis, “Distributed estimation algorithms for non-linear systems,” *Automatic Control, IEEE Transactions on*, vol. 30, no. 5, pp. 418–425, 1985.
- [43] P. Varshney *et al.*, “Distributed Bayesian hypothesis testing with distributed data fusion,” *Systems, Man and Cybernetics, IEEE Transactions on*, vol. 18, no. 5, pp. 695–699, 1988.
- [44] J. A. Gubner, “Distributed estimation and quantization,” *Information Theory, IEEE Transactions on*, vol. 39, no. 4, pp. 1456–1459, 1993.
- [45] A. Ribeiro and G. Giannakis, “Bandwidth-constrained distributed estimation for wireless sensor networks-part I: Gaussian case,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 3, pp. 1131–1143, March.
- [46] —, “Bandwidth-constrained distributed estimation for wireless sensor networks-part II: unknown probability density function,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 7, pp. 2784–2796, July 2006.
- [47] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, “Joint estimation in sensor networks under energy constraints,” in *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, Oct. 2004, pp. 264–271.
- [48] J.-J. Xiao and Z.-Q. Luo, “Decentralized estimation in an inhomogeneous sensing environment,” *Information Theory, IEEE Transactions on*, vol. 51, no. 10, pp. 3564–3575, 2005.
- [49] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, “Power scheduling of universal decentralized estimation in sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 2, pp. 413–422, Feb. 2006.
- [50] J.-J. Xiao, A. Ribeiro, Z.-Q. Luo, and G. B. Giannakis, “Distributed compression-estimation using wireless sensor networks,” *Signal Processing*

Magazine, IEEE, vol. 23, no. 4, pp. 27–41, 2006.

- [51] J.-J. Xiao, S. Cui, Z.-Q. Luo, and A. Goldsmith, “Linear coherent decentralized estimation,” *Signal Processing, IEEE Transactions on*, vol. 56, no. 2, pp. 757–770, Feb. 2008.
- [52] G. Barrenetxea, F. Ingelrest, G. Schaefer, and M. Vetterli, “The hitchhiker’s guide to successful wireless sensor network deployments,” in *SenSys ’08: Proceedings of the 6th ACM conference on Embedded network sensor systems*. New York, NY, USA: ACM, 2008, pp. 43–56.
- [53] T. W. Hnat, V. Srinivasan, J. Lu, T. I. Sookoor, R. Dawson, J. Stankovic, and K. Whitehouse, “The hitchhiker’s guide to successful residential sensing deployments,” in *Proceedings of the 9th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2011, pp. 232–245.
- [54] E. M. Shakshuki, H. Malik, and T. R. Sheltami, “Lessons learned: Simulation vs wsn deployment,” in *Proc. International Conference on Advanced Information Networking and Applications AINA ’09*, 2009, pp. 580–587.
- [55] R. Ghez, *Diffusion Phenomena*. New York: Kluwer Academic/Plenum Publishers, 2001.
- [56] V. Turchin and A. Gershman, “Sensor array approach to nonwave field processing,” in *Acoustics, Speech, and Signal Processing, 1995. ICASSP-95., 1995 International Conference on*, vol. 3, May 1995, pp. 2072–2075.
- [57] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 2: Detection Theory*. Prentice Hall, 1998.
- [58] S. Lasassmeh and J. Conrad, “Time synchronization in wireless sensor networks: A survey,” in *IEEE SoutheastCon 2010 (SoutheastCon), Proceedings of the*, 2010, pp. 242–245.
- [59] N. Patwari, J. Ash, S. Kyperountas, A. Hero, R. Moses, and N. Correal, “Locating the nodes: cooperative localization in wireless sensor networks,” *Signal Processing Magazine, IEEE*, vol. 22, no. 4, pp. 54–69, 2005.
- [60] W. Bajwa, “Matched source-channel communication for field estimation in wireless sensor networks,” in *Proc. the Fourth Int. Symposium on Information Processing in Sensor Networks*, 2005, pp. 332–339.
- [61] D. Stoyan, *Stochastic geometry and its applications*. Chichester New York: Wiley, 1995.

- [62] Z. Chair and P. K. Varshney, "Optimal data fusion in multiple sensor detection systems," *IEEE Transactions On Aerospace And Electronic Systems*, vol. AES-22, no. 1, pp. 98–101, 1986.
- [63] S. Aldalahmeh, M. Ghogho, and A. Swami, "Fast distributed detection, localization, and estimation of a diffusive target in wireless sensor networks," in *Wireless Communication Systems, 2010. ISWCS2010, Seventh International Symposium on*, York, Sept. 2010.
- [64] S. Aldalahmeh and M. Ghogho, "Robust distributed detection, localization, and estimation of a diffusive target in clustered wireless sensor networks," in *International Conference on Acoustics Speech and Signal Processing (ICASSP) 2011*. Prague, Czech Republic: IEEE, May 2011.
- [65] S. A. Aldalahmeh, M. Ghogho, and A. Swami, "Distributed detection of an unknown target in clustered wireless sensor networks," in *2011 IEEE 12th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2011)*, San Francisco, USA, Jun. 2011.
- [66] S. A. Aldalahmeh and M. Ghogho, "Statistical analysis of optimal distributed detection fusion rule in wireless sensor networks," in *Wireless Advanced 2012 (WiAd 2012)*, London, United Kingdom, Jun. 2012.
- [67] —, "Hard decision fusion in censored wireless sensor networks over rayleigh fading multiple access channel," *Signal Processing, IEEE Transactions on*, submitted.
- [68] M. G. Sami A Aldalahmeh and D. McLernon, "Contention delay analysis of multipacket reception slotted aloha in censored wireless sensor networks," *Vehicle Technology, IEEE Transactions on*, submitted.
- [69] J. G. Andrews, R. K. Ganti, M. Haenggi, N. Jindal, and S. Weber, "A Primer on Spatial Modeling and Analysis in Wireless Networks," *IEEE Communications Magazine*, vol. 48, no. 11, pp. 156–163, Nov. 2010, available at
- [70] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic Geometry and Random Graphs for the Analysis and Design of Wireless Networks," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009, invited Paper. 2010 IEEE Communication Society Best Tutorial Paper Award. Available at
- [71] R. Streit, *Poisson point processes imaging, tracking, and sensing*. New York:

Springer, 2010.

- [72] S. Thomopoulos, R. Viswanathan, and D. Bougoulas, "Optimal decision fusion in multiple sensor systems," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. AES-23, no. 5, pp. 644–653, sept. 1987.
- [73] M. Guerriero, P. Willett, and J. Glaz, "Distributed target detection in sensor networks using scan statistics," *Signal Processing, IEEE Transactions on*, vol. 57, no. 7, pp. 2629–2639, july 2009.
- [74] M. Guerriero, L. Svensson, and P. Willett, "Bayesian data fusion for distributed target detection in sensor networks," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3417–3421, June 2010.
- [75] B. Chen, L. Tong, and P. Varshney, "Channel-aware distributed detection in wireless sensor networks," *Signal Processing Magazine, IEEE*, vol. 23, no. 4, pp. 16–26, July 2006.
- [76] 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, Nov. 2005.
- [77] M. Gastpar, B. Rimoldi, and M. Vetterli, "To code, or not to code: lossy source-channel communication revisited," *Information Theory, IEEE Transactions on*, vol. 49, no. 5, pp. 1147–1158, May 2003.
- [78] B. Liu and B. Chen, "Type-based decentralized detection in wireless sensor networks," in *Proc. 2006 Int. Conf. on Signal Process*, vol. 55, pp. 1899–1910, 2005.
- [79] A. Anandkumar and L. Tong, "Type-based random access for distributed detection over multiaccess fading channels," *Signal Processing, IEEE Transactions on*, vol. 55, no. 10, pp. 5032–5043, oct. 2007.
- [80] F. Li and J. S. Evans, "Decision fusion over noncoherent fading multiaccess channels," *Global Telecommunications Conference 2008 IEEE GLOBECOM 2008 IEEE*, pp. 1–5, 2008.
- [81] M. Banavar, A. Smith, C. Tepedelenlioglu, and A. Spanias, "Distributed detection over fading macs with multiple antennas at the fusion center," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, march 2010, pp. 2894–2897.
- [82] C. R. Berger, M. Guerriero, S. Zhou, and P. Willett, "Pac vs. mac for decentralized detection using noncoherent modulation," *Trans. Sig. Proc.*, vol. 57,

- no. 9, pp. 3562–3575, Sep. 2009.
- [83] M. Vu and A. Paulraj, “Mimo wireless linear precoding,” *Signal Processing Magazine, IEEE*, vol. 24, no. 5, pp. 86–105, 2007.
- [84] F. Li, J. Evans, and S. Dey, “Decision fusion over noncoherent fading multi-access channels,” *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4367–4380, sept. 2011.
- [85] H. Budzier, *Thermal infrared sensors theory, optimization, and practice*. Chichester, West Sussex, U.K. Hoboken, N.J: Wiley, 2011.
- [86] F. Digham, M. Alouini, and M. Simon, “On the energy detection of unknown signals over fading channels,” in *Communications, 2003. ICC '03. IEEE International Conference on*, vol. 5, May 2003, pp. 3575 – 3579 vol.5.
- [87] R. Tan, G. Xing, B. Liu, and J. Wang, “Impact of data fusion on real-time detection in sensor networks,” *Real-Time Systems Symposium, IEEE International*, vol. 0, pp. 323–332, 2009.
- [88] X. Zhang, H. Poor, and M. Chiang, “Power allocation in distributed detection with wireless sensor networks,” in *Military Communications Conference, 2006. MILCOM 2006. IEEE*, Oct. 2006, pp. 1–7.
- [89] 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, Dec. 2004.
- [90] A. Nehorai, B. Porat, and E. Paldi, “Detection and localization of vapor-emitting sources,” *Signal Processing, IEEE Transactions on*, vol. 43, no. 1, pp. 243–253, Jan 1995.
- [91] P. Stoica and N. Arye, “Music, maximum likelihood, and cramer-rao bound,” *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 37, no. 5, pp. 720–741, May 1989.
- [92] A. Jeremic and A. Nehorai, “Landmine detection and localization using chemical sensor array processing,” *Signal Processing, IEEE Transactions on*, vol. 48, no. 5, pp. 1295–1305, May 2000.
- [93] —, “Design of chemical sensor arrays for monitoring disposal sites on the ocean floor,” *Oceanic Engineering, IEEE Journal of*, vol. 23, no. 4, pp. 334–343, Oct. 1998.
- [94] T. Zhao and A. Nehorai, “Detecting and estimating biochemical dispersion of

- a moving source in a semi-infinite medium,” *Signal Processing, IEEE Transactions on*, vol. 54, no. 6, pp. 2213–2225, June 2006.
- [95] S. Vijayakumaran, Y. Levinbook, and T. F. Wong, “Maximum likelihood localization of a diffusive point source using binary observations,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 2, pp. 665–676, 2007.
- [96] J. Matthes, L. Groll, and H. Keller, “Source localization by spatially distributed electronic noses for advection and diffusion,” *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1711 – 1719, May 2005.
- [97] M. Ortner, A. Nehorai, and A. Jerémic, “Biochemical transport modeling and bayesian source estimation in realistic environments,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 6, pp. 2520–2532, Jun. 2007.
- [98] M. Ortner and A. Nehorai, “A sequential detector for biochemical release in realistic environments,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 8, pp. 4173–4182, Aug. 2007.
- [99] B. Porat and A. Nehorai, “Localizing vapor-emitting sources by moving sensors,” *Signal Processing, IEEE Transactions on*, vol. 44, no. 4, pp. 1018–1021, Apr. 1996.
- [100] P. Tzanos, M. Zefran, and A. Nehorai, “Information based distributed control for biochemical source detection and localization,” in *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, Apr. 2005, pp. 4457 – 4462.
- [101] Y. Wang, R. Tan, G. Xing, J. Wang, and X. Tan, “Accuracy-aware aquatic diffusion process profiling using robotic sensor networks,” in *Proceedings of the 11th international conference on Information Processing in Sensor Networks*, ser. IPSN '12. New York, NY, USA: ACM, 2012, pp. 281–292.
- [102] T. Zhao and A. Nehorai, “Information-driven distributed maximum likelihood estimation based on gauss-newton method in wireless sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 9, pp. 4669–4682, 2007.
- [103] S. Zejnilovic, J. Gomes, and B. Sinopoli, “Collaborative diffusive source localization in wireless sensor networks,” in *Signal Processing Conference (EU-SIPCO), 2012 Proceedings of the 20th European*, Aug. 2012, pp. 704–708.
- [104] T. Zhao and A. Nehorai, “Distributed sequential Bayesian estimation of a diffusive source in wireless sensor networks,” *Signal Processing, IEEE Transactions on*, vol. 55, no. 4, pp. 1511–1524, April 2007.

- [105] L. Rossi, B. Krishnamachari, and C.-C. Kuo, "Distributed parameter estimation for monitoring diffusion phenomena using physical models," in *Sensor and Ad Hoc Communications and Networks, 2004. IEEE SECON 2004. 2004 First Annual IEEE Communications Society Conference on*, Oct. 2004, pp. 460–469.
- [106] L. A. Rossi, B. Krishnamachari, and C.-C. J. Kuo, "Monitoring of diffusion processes with pde models in wireless sensor networks," in *in Defense and Security Symposium 2004*, 2004, pp. 12–16.
- [107] E. Fox, J. Fisher, and A. Willsky, "Detection and localization of material releases with sparse sensor configurations," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1886–1898, May 2007.
- [108] J. M. Stockie, "The mathematics of atmospheric dispersion modeling," *Siam Review*, vol. 53, no. 2, pp. 349–372, 2011.
- [109] F. Sawo, K. Roberts, and U. D. Hanebeck, "Bayesian estimation of distributed phenomena using discretized representations of partial differential equations," in *3rd International Conference on Informatics in Control, Automation and Robotics (ICINCO)*, 2006, pp. 16–23.
- [110] M. F. Huber and U. D. Hanebeck, "The hybrid density filter for nonlinear estimation based on hybrid conditional density approximation," in *Information Fusion, 2007 10th International Conference on*. IEEE, 2007, pp. 1–8.
- [111] F. Sawo, M. F. Huber, and U. D. Hanebeck, "Parameter identification and reconstruction for distributed phenomena based on hybrid density filter," in *Information Fusion, 2007 10th International Conference on*. IEEE, 2007, pp. 1–8.
- [112] V. Klumpp, F. Sawo, U. Hanebeck, and D. Franken, "The sliced gaussian mixture filter for efficient nonlinear estimation," in *Information Fusion, 2008 11th International Conference on*, 30 2008–July 3 2008, pp. 1–8.
- [113] F. Sawo, V. Klumpp, and U. Hanebeck, "Simultaneous state and parameter estimation of distributed-parameter physical systems based on sliced gaussian mixture filter," in *Information Fusion, 2008 11th International Conference on*, 30 2008–July 3 2008, pp. 1–8.
- [114] S. Sundhar and V. Veeravalli, "Localization and intensity tracking of diffusing point sources using sensor networks," in *Global Telecommunications Conference, 2007. GLOBECOM'07. IEEE*. IEEE, 2007, pp. 3107–3111.

- [115] S. Sundhar Ram, V. Veeravalli, and A. Nedic, "Incremental recursive prediction error algorithm for parameter estimation in sensor networks," in *Information Fusion, 2008 11th International Conference on*. IEEE, 2008, pp. 1–8.
- [116] I. Dokmanic, J. Ranieri, A. Chebira, and M. Vetterli, "Sensor networks for diffusion fields: detection of sources in space and time," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*. IEEE, 2011, pp. 1552–1558.
- [117] P. L. Dragotti, M. Vetterli, and T. Blu, "Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets Strang–fix," *Signal Processing, IEEE Transactions on*, vol. 55, no. 5, pp. 1741–1757, 2007.
- [118] J. Ranieri, I. Dokmanic, A. Chebira, and M. Vetterli, "Sampling and reconstruction of time-varying atmospheric emissions," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3673–3676.
- [119] R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter estimation and inverse problems*. Academic Press, 2012.
- [120] P. Kathirgamanathan, R. McKibbin, and R. I. McLachlan, "Source release-rate estimation of atmospheric pollution from a non-steady point source at a known location," *Environmental Modeling and Assessment*, vol. 9, no. 1, pp. 33–42, 2004.
- [121] P. Kathirgamanathan, R. McKibbin, and R. McLachlan, "Source release rate estimation of atmospheric pollution from a non-steady point source—part 2: Source at an unknown location," *Research Letters in the Information and Mathematical Sciences*, vol. 5, pp. 85–118, 2003.
- [122] H. Chen and X. R. Li, "Joint identification and tracking of multiple cbrne clouds based on sparsity pursuit," in *Information Fusion (FUSION), 2010 13th Conference on*. IEEE, 2010, pp. 1–8.
- [123] J. C. Chen, K. Yao, and R. E. Hudson, "Source localization and beamforming," *Signal Processing Magazine, IEEE*, vol. 19, no. 2, pp. 30–39, 2002.
- [124] A. Haji-Sheikh, *Heat Conduction Using Green's Functions*. Washington: Hemisphere Pub. Corp, 1992.
- [125] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory (v. 1)*. Prentice Hall, 1993.

- [126] G. Golub and V. Pereyra, "Separable nonlinear least squares: the variable projection method and its applications," *Inverse Problems*, vol. 19, no. 2, p. R1, 2003. [Online]. Available:
- [127] B. Zhang, X. Cheng, N. Zhang, Y. Cui, Y. Li, and Q. Liang, "Sparse target counting and localization in sensor networks based on compressive sensing," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 2255–2263.
- [128] S. Boyd, *Convex Optimization*. Cambridge England: Cambridge, 2004.
- [129] Y. Yang and T.-S. Yum, "Delay distributions of slotted Aloha and CSMA," *Communications, IEEE Transactions on*, vol. 51, no. 11, pp. 1846 – 1857, nov. 2003.
- [130] Y. Wang, M. Vuran, and S. Goddard, "Cross-layer analysis of the end-to-end delay distribution in wireless sensor networks," *Networking, IEEE/ACM Transactions on*, vol. 20, no. 1, pp. 305 –318, feb. 2012.
- [131] L. Liu and J.-F. Chamberland, "Cross-layer optimization and information assurance in decentralized detection over wireless sensor networks," in *Signals, Systems and Computers, 2006. ACSSC '06. Fortieth Asilomar Conference on*, 29 2006-nov. 1 2006, pp. 271 –275.
- [132] Y. Wang, M. Vuran, and S. Goddard, "Analysis of event detection delay in wireless sensor networks," in *INFOCOM, 2011 Proceedings IEEE*, april 2011, pp. 1296 –1304.
- [133] V. Naware, G. Mergen, and L. Tong, "Stability and delay of finite-user slotted aloha with multipacket reception," *Information Theory, IEEE Transactions on*, vol. 51, no. 7, pp. 2636–2656, 2005.
- [134] S. Sen, D. Dorsey, R. Guerin, and M. Chiang, "Analysis of slotted aloha with multipacket messages in clustered surveillance networks," in *MILITARY COMMUNICATIONS CONFERENCE, 2012 - MILCOM 2012*, 2012, pp. 1–6.
- [135] P. Levis, N. Lee, M. Welsh, and D. Culler, "Tossim: accurate and scalable simulation of entire tinyos applications," in *Proceedings of the 1st international conference on Embedded networked sensor systems*, ser. SenSys '03. New York, NY, USA: ACM, 2003, pp. 126–137.
- [136] M. Osborne, "Separable least squares, variable projection, and the gauss-newton algorithm." *ETNA. Electronic Transactions on Numerical Analysis [electronic only]*, vol. 28, pp. 1–15, 2007.