# Collaborative Information Processing in Wireless Sensor Networks for Diffusive Source Estimation

by

**Kamrul Hakim**

PH.D. DISSERTATION

The University of New Mexico

Albuquerque, New Mexico

March, 2013

Submitted by: Kamrul Hakim

Supervisor: Dr. Sudharman K. Jayaweera
Department of Electrical and Computer Engineering
University of New Mexico

Committee Members: Dr. Nasir Ghani
Department of Electrical and Computer Engineering
University of New Mexico

Dr. Patrick Bridges
Department of Computer Science
University of New Mexico

Dr. Khanh D. Pham
Air Force Research Laboratory (AFRL)
Albuquerque, NM

# Dedication

*To my loving parents*

*"Read! In the name of your Lord who has created. He has created man from a clot. Read! And your Lord is the most Generous. Who has taught by the pen. He has taught man that which he knew not"* -Al Quran, 96:1-5

# Acknowledgements

I would like to express my sincere gratitude to my academic supervisor, Professor Sudharman Jayaweera for his invaluable guidance and support during my graduate study in Communication and Information Sciences Laboratory (CISL) at University of New Mexico (UNM). I would like to thank him for all his insights, thought provoking questions and guidance throughout my Ph.D. study. I am also thankful to Dr. Nasir Ghani, Dr. Patrick Bridges and Dr. Khanh Pham to agreeing to serve as members in my dissertation committee even at hardship.

I would like to thank all my Bangladeshi friends, and also my colleagues at CISL with whom I have had the pleasure of working over the years. My stay at UNM was made pleasant and enjoyable because of their friendly support. I also take this opportunity to thank administrative staff at ECE for their constant support and help.

Mostly, this work would not be possible without love, encouragement and support of my family. I am grateful to my parents Abdul Hakim and Selina Hakim who had been a constant source of love and inspiration in all my pursuits. Their prayer for me is what sustained me thus far. I would like to thank my loving wife Monami for her love, understanding and encouragement throughout this experience. I also would like to thank my dearest siblings, Sunny and Tani, and my grandmother for always standing by my side.

# Collaborative Information Processing in Wireless Sensor Networks for Diffusive Source Estimation

by

**Kamrul Hakim**

Ph.D., Electrical and Computer Engineering, University of New Mexico, 2013

## Abstract

In this dissertation, we address the issue of collaborative information processing for diffusive source parameter estimation using wireless sensor networks (WSNs) capable of sensing in dispersive medium/environment, from signal processing perspective.

We begin the dissertation by focusing on the mathematical formulation of a special diffusion phenomenon, i.e., an underwater oil spill, along with statistical algorithms for meaningful analysis of sensor data leading to efficient estimation of desired parameters of interest. We first derive physical model for an underwater oil spill scenario to obtain the spatio-temporal concentration distribution of the dispersed substance. Our motivation is to obtain an analytical solution to the problem, rather than using non-model based sophisticated numerical techniques. We tried to make the physical diffusion model as much appropriate as possible, while maintaining some pragmatic and reasonable assumptions for the simplicity of exposition and analytical derivation.

The dissertation studies both source localization and tracking for static and moving diffusive sources respectively. For static diffusive source localization, we investigate two

parametric estimation techniques based on the maximum-likelihood (ML) and the best linear unbiased estimator (BLUE) for a special case of our obtained physical dispersion model. We prove the consistency and asymptotic normality of the obtained ML solution when the number of sensor nodes and samples approach infinity, and derive the Cramer-Rao lower bound (CRLB) on its performance. In case of a moving diffusive source, we propose a particle filter (PF) based target tracking scheme for moving diffusive source, and analytically derive the posterior Cramer-Rao lower bound (PCRLB) for the moving source state estimates as a theoretical performance bound.

Further, we explore nonparametric or machine learning based estimation technique for diffusive source parameter estimation using Dirichlet process mixture model (DPMM). Since real data are often complicated, no parametric model is suitable. As an alternative, we exploit the rich tools of nonparametric Bayesian methods, in particular the DPMM, which provides us with a flexible and data-driven estimation process. We propose DPMM based static diffusive source localization algorithm and provide analytical proof of convergence. The proposed algorithm is also extended to the scenario when multiple diffusive sources of same kind are present in the diffusive field of interest.

Efficient power allocation can play an important role in extending the lifetime of a resource constrained WSN. Resource-constrained WSNs rely on collaborative signal and information processing for efficient handling of large volumes of data collected by the sensor nodes. In this dissertation, the problem of collaborative information processing for sequential parameter estimation in a WSN is formulated in a cooperative game-theoretic framework, which addresses the issue of fair resource allocation for estimation task at the Fusion center (FC). The framework allows addressing either resource allocation or commitment for information processing as solutions of cooperative games with underlying theoretical justifications. Different solution concepts found in cooperative games, namely, the Shapley function and Nash bargaining are used to enforce certain kinds of fairness among the nodes in a WSN.

# Contents

# List of Figures

# Glossary

| | |
|---|---|
| AWGN | Additive white Gaussian Noise |
| BLUE | Best linear unbiased estimator |
| CDMA | Code division multiple access |
| C-DSL | Centralized dynamic spectrum leasing |
| CRLB | Cramér-Rao lower bound |
| CSI | Channel state information |
| DPMM | Dirichlet process mixture model |
| DSA | Dynamic spectrum access |
| DSL | Dynamic spectrum leasing |
| DSS | Dynamic spectrum sharing |
| FC | Fusion center |
| FIM | Fisher information matrix |
| H-SVM | Hierarchical SVM |
| IC | Interference cap |

| | |
|---|---|
| LMMSE | Linear minimum mean-squared error |
| MEMS | Micro-electro-mechanical systems |
| MF | Matched filter |
| ML | Maximum-likelihood |
| MLE | Maximum-likelihood estimator |
| MPA | Maximum power allocation |
| MSE | Mean squared error |
| MUD | Multiuser detector |
| NB | Nash bargaining |
| NBS | Nash bargaining solution |
| NE | Nash equilibrium |
| PCRLB | Posterior Cramér-Rao lower bound |
| PDF | Probability density function |
| PF | Particle filter |
| PoI | Phenomenon of Interest |
| POMDP | Partially observable Markov decision process |
| QoS | Quality-of-Service |
| RL | Reinforcement learning |
| RMSE | Root mean squared error |

| | |
|---|---|
| SINR | Signal-to-interference-plus-noise ratio |
| SNR | Signal-to-noise ratio |
| SLT | Statistical learning theory |
| SVM | Support vector machine |
| UAV | Unmanned autonomous vehicles |
| WSN | Wireless sensor networks |

# Chapter 1

# Motivations and Overview of the Dissertation

## 1.1 Introduction

The release of hazardous diffusive substance into environment, i.e., the release of chemicals from industries into air or liquid petroleum hydrocarbon into the ocean or coastal water due to human activity, has attracted tremendous attention because of its environmental, biological and economical impact. Recent *BP oil disaster* in the Gulf of Mexico is a perfect example of how spill stemmed from a sea-floor oil gusher can severely damage the marine and wildlife habitats as well as the Gulf's fishing and tourism industries [1, 2]. Hence, research is required to model such scenarios to assist in planning and emergency decision making, thereby reducing the threats and hazardous effect on environment and economy. In order to limit such potential damage, it is important to have an efficient and reliable detection and estimation system to predict substance dispersion. To that end, over the last decades, researchers and scientists in the signal processing community have given considerable amount of attention in addressing the difficult task of diffusive source estimation, i.e., diffusive source

localization [3–6]. Research in diffusive source parameter estimation can, in general, be extended to many other similar contexts such as homeland security, environmental and industrial monitoring, pollution control, servers and data center temperature monitoring as well [3, 6–12]. Some useful applications of diffusive source estimation include, among many others, detection of potential biochemical attacks, sensing explosives mounted on a static platform or a moving vehicle, and detecting leakage of hazardous liquids from a container. In this dissertation, we address the issue of collaborative information processing for diffusive source parameter estimation from signal processing perspective to improve on and overcome the limitations in the methods found in existing literature, and then further develop novel and efficient methods for solving realistic diffusion scenarios.

Wireless sensor networks, in general, are tasked with information collection, processing and dissemination in dynamic and hazardous environments. Sensor networks are prone to premature failure since nodes might run out of their batteries rapidly due to work load variations (e.g. relay nodes or nodes that are in close proximity to a certain PoI are triggered frequently), different communication environments or hardware setup. It is undesirable for a sensor node to waste power as excessive use of battery power can shorten the lifetime of a node. Hence, a commonly used performance criteria for many wireless sensor networks is the network lifetime while satisfying required coverage and connectivity over the deployment region. In this dissertation, we address the challenge of *fair* resource allocation, i.e. power, among sensing nodes in a resource constrained WSN tasked with parameter estimation, by developing algorithms based on cooperative game-theoretic framework for collaborative sensor signal processing.

## 1.2 Recent Trends in Wireless Sensor Networks

Wireless sensor networks (WSNs) are spatially distributed data acquisition systems consisting of many sensing nodes tasked with cooperatively monitoring a Phenomenon of Interest

(PoI) [13, 14]. Sensor nodes are capable of a limited amount of local processing and wireless communication. Though having limited capability of processing information, when large number of these sensors communicate and share information among themselves, they can measure a desired PoI in great detail. Figure 1.1 shows a typical WSN connected to a IP (internet protocol) network

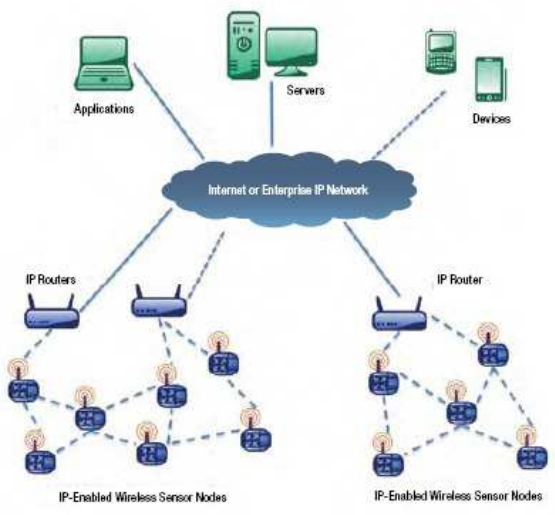



Figure 1.1: A typical wireless sensor network; [http://rtcmagazine.com/]

WSNs have applications in a variety of fields because of their inexpensive feature and ad-hoc method of deployment. In order to collect data from the diffusive field and to operate on those obtained data to extract useful information about the diffusive source and the field itself, we propose to use WSNs. The reasons behind choosing WSNs in our study are *distributed and ad-hoc nature of deployment*, and *wireless sensing and communication capability*. For example, when exact location of a PoI is not known, distributed sensing would allow closer placement to the phenomenon than a single sensor would permit. Also, in many practical cases, network of sensors are required to avoid possible human deployment because of hazardous environment, or to overcome environmental obstacles such as obstructions, line of sight constraints etc. In situations, where no fixed infrastructure for power or

communication are available, network of sensors are preferable to single sensing node to cope up with finite sources of energy and communications. Another important benefit of using WSNs is the distributed wireless sensing, processing and communication capability. Unlike a centralized system, distributed sensor nodes in a WSN can sometimes process obtained data locally as much as possible in order to minimize the total number of bits to be transmitted to the fusion center (FC) which leads to savings in overall energy consumption.

Recent advances in low-power micro sensors, integrated circuits and wireless communication technologies have enabled the development of WSNs for variety of useful applications [13, 15, 16]. Also with the developments of unmanned autonomous vehicles (UAV's), WSNs are gaining popularity due to their potential to be useful for a wide range of applications including environmental monitoring, intrusion detection, and various military and civilian applications [13,17,18]. Due to advanced micro-electro-mechanical systems (MEMS), many of the state-of-the-art sensors are now more accurate, robust against noise and energy efficient [19–22]. These new cutting-edge sensors can withstand severe unfavorable conditions in hazardous areas where human deployment is impossible. Modern sensor nodes can now be deployed either in a systematic or in random pattern in various environments, such as, on ground, in the air, under water, on vehicle bodies depending on feasibilities and requirements.

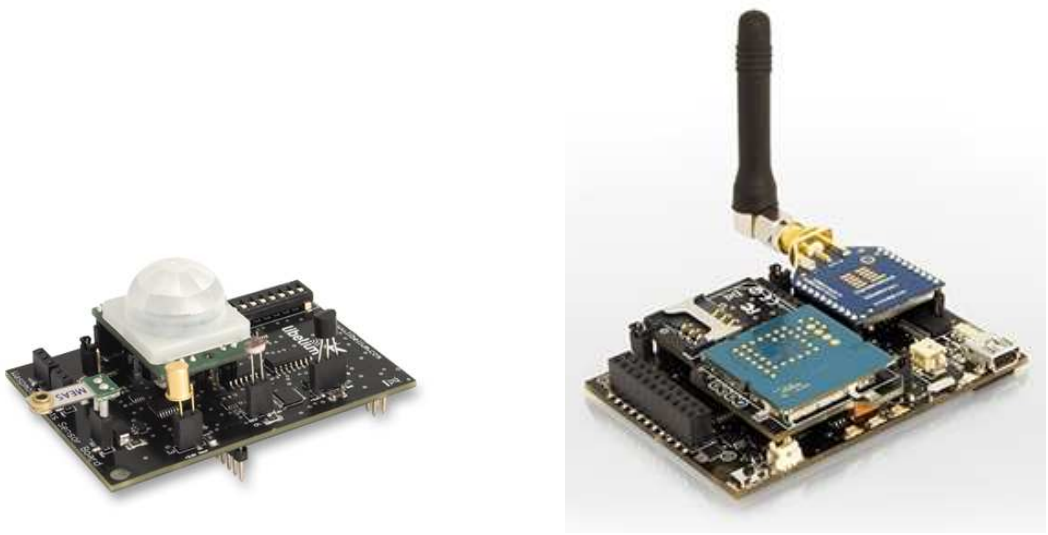
For instance, KMHP-100 MEMS microhotplate (in Fig. 1.2(a)) suitable for chemical sensor applications can operate continuously at up to $650^oC$ ensuring stable operation in hostile environment [20]. Some of the important commercially available cheap chemical sensors which are useful in sensing gas diffusion are SeaPORT SC-210 and EX-TEC OD 4 [21,23]. SeaPORT SC-210 sensor developed by Seacoast Science is a chemical sensor which can accurately detect airborne organic chemicals (in Fig. 1.2(b)). EX-TEC OD 4 is an odorant measuring sensor used for ensuring safety in natural gas supply. Among underwater sensors, LI-192SA underwater quantum sensor can accurately measures photosynthetically active radiation (PAR) in freshwater or saltwater environments [22]. The LI-193 sensor fea-


(a) (b)

Figure 1.2: Typical chemical and gas sensors: (a) KMHP-100 chemical sensor [http://www.kebaili.com], (b) SeaPORT SC-210 gas sensor. [http://www.technologyreview.com]

tures a high sensitivity optical design and can be used at pressures up to 493 psi (pound per square inch) or 350 meter depths. Recently researchers in Sandia National laboratory invented micro-chemical sensors for *in-situ* monitoring and characterization of volatile contaminants for real-time monitoring of subsurface contaminants [19]. These SWNT TFT (Single-Walled Carbon Nanotubes Thin Film Transistor) sensors are well-suited for detecting chemical weapons agents and explosives, which typically occur in low concentrations *in situ*. Some of the commercially available but cheap gas sensors are: Safety Siren Pro HS80504 sensor (for detection of carbon monoxide gas), Amprobe GSD600 gas detector for Methane and Propane, TPI A739 combustible gas detector, and GE Security 60-652-95 -ITI wireless carbon monoxide (CO) sensor [24–26].

A typical sensing unit of a sensor node is equipped with a programmable processor board with RF transceiver for wireless communication. Figure 1.3 shows a Waspmote by Libelium



(a) (b)

Figure 1.3: (a) Sensing board, and (b) RF transceiver of a typical sensor node with wireless communication capability [source: http://www.libelium.com].

with gas sensing and transceiver units. Waspmote by Libelium is equipped with GPRS, Bluetooth, Wifi, and RFID/NFC modules [27].

All of these aforementioned exciting cutting-edge features in recently developed sensors make them suitable candidates for the set of applications involving monitoring of diffusion phenomena that we are interested in.

## 1.3 Limitations of Existing Approaches

Source or target localization using distributed sensor arrays is an area of active research interest for a considerable period of time [28, 29]. In the past, detection and localization problems of diffusive sources in WSN have been a topic of interest, specially in the case of chemical/biological threat detection. Interesting research in this context can be found in [3, 5, 6, 30], where biochemical concentration distribution in space and time for different

types of diffusive sources, diffusion models and/or sensor networks are estimated. For instance, remotely localizing a gas or odor source using mobile robot was proposed in [6] by fitting the gas distribution model to the gas sensor response at the sensor locations. However, the mobile sensor dynamics model therein, was obtained empirically which does not allow for dynamic environment and moving diffusive source. In [3], a maximum likelihood (ML) estimator was developed for localizing vapor emitting sources and its asymptotic normality of the obtained maximum likelihood (ML) estimator was proved when the signal-to-noise-ratio (SNR) approaches infinity. Many other estimation techniques have also been used in diffusive source parameters estimation literature [4, 5, 30–32]. In particular, Bayesian estimation has been applied in [30, 31] in a sequential manner, which is not suitable in many practical scenarios where faster estimation and immediate actions based on the estimation are top priorities. A real-time maximum-likelihood estimation method was proposed in [4] for estimating diffusive source parameters, where consistency and asymptotic efficiency of the obtained estimator were proved when the density of sensors becomes infinite. In [33], the problem of impulsive diffusive source localization was solved assuming the spatial sensor measurements at any sensor location as a scaled and shifted version of a common prototype function, leading to solving a set of linear equations. However, the physical diffusion models used in [4, 33] are oversimplified with the diffusive sources assumed to be impulsive or instantaneous in nature.

Although research has been done in tracking and/or estimating time-varying parameter estimation in general [34–38], very few attempts have been made in time varying diffusive source parameter estimation. Some of these methods cannot be applied directly into our time varying parameter estimation model since, e.g., for a moving source, the concentration at the current time is affected by all past values of source position. Therefore, time-cumulation effects on the concentrations (i.e. observations) must be taken into account to estimate time-varying parameters. Among previous works, a parametric moving path model for a diffusive moving source was discussed in [5], where the moving source path was approximated using finite number basis functions. Tracking performance in this case depends on the smoothness

of the source trajectory, prior information about the moving source trajectory and choosing a suitable finite set of basis functions. In [39], a novel recursive algorithm was proposed to track the intensity of a diffusive point source, but the source location was considered as an unknown static value. The aforementioned limitations may be overcome by developing or exploiting state-of-the-art Bayesian-based location tracking methods, i.e., Kalman filter [40], particle filter (PF) [41] suitable for dynamic state estimation problem.

There has been flurry of research in diffusive source detection and localization using various aforementioned parametric estimation techniques. However, to the best of our knowledge, no attempts have been made in exploiting the rich collection of tools of machine learning for diffusive source parameter estimation and localization. Though parametric models provides us with a rigorous analysis of many fundamental questions for inference, if instead data is sparse and prior knowledge is limited, robust nonparametric methods for inference are generally preferred [42–44]. Since real data are often complicated, no parametric model is suitable sometimes. This leads to degradation of the accuracy of parametric estimation techniques. Among recent works, support vector machine (SVM) and hierarchical-SVM (H-SVM) were successfully applied in acoustic objection localization, and estimating physical location of sensor nodes and channel noise in [45], [46] respectively with theoretical and experimental validations. The problem of kernel parameter estimation for manifold learning is addressed based on the nonparametric statistical theory estimation was proposed in [47]. The authors in [48] proposed a novel statistical learning theory (SLT) based model selection approach to estimate optimal motion models from small data sets of image measurements for single motion estimation and tracking. In [49], reinforcement learning (RL) method has been used for parameter estimation in statistical spoken dialogue systems modeled as a partially observable Markov decision process (POMDP).

For efficient resource allocation problem among sensing nodes, cooperative game theoretic concepts can be a useful tool in approaching collaborative signal processing problems in distributed sensor networks. Unlike non-cooperative game theory, where individual players

compete with each other to achieve their goals of maximizing individual payoffs, cooperative game theory allows competing players (or nodes) to form coalitions so as to efficiently achieve their individual goals. Although collaborative signal and information processing for sensor networks have been extensively studied in the literature [35–37, 50–52], there is still ample room for research aimed at developing a formal analytical framework for collaborative information processing in resource-constrained sensor networks, especially when maximizing the sensor network lifetime is the most important objective. The fact that nodes act to optimize their own payoffs has triggered research on using non-cooperative game theory to manage powers in wireless sensor networks [53–55]. On the other hand, for a resource constrained WSN, cooperative game theory can be a natural choice and comes in handy when estimating a parameter with desired estimator quality is the ultimate goal. In current literature, very few attempts have been made to exploit the rich collection of cooperative game theory in power/energy-constrained WSNs tasked with estimating a parameter. For example, in [56], a novel concept of incompletely cooperative game theory was used to simultaneously achieve energy conservation and throughput for WSNs. On the other hand, [57, 58] used cooperative game theory for channel/bandwidth allocation problem.

## 1.4 Main Contributions

In this section, we outline the significant contributions presented in the dissertation:

- We formulate the problem of an underwater oil spill scenario, and derive mathematical model for the space-time substance dispersion considering the effect of laminar water velocity as an external force. For a special case of our obtained physical dispersion model, we propose and implement maximum-likelihood (ML) and best linear unbiased estimator (BLUE) based parameter estimation techniques for a static diffusive source continuously emitting substance. We prove both the consistency and asymptotical normality of our obtained ML based solution when the number of sensor nodes and

time samples go to infinity. Derivation of the Cramér-Rao lower bound (CRLB) as a theoretical performance bound is also presented.

- For moving diffusive source emitting substance into dispersive medium, we propose a particle filter (PF) based target tracking method. To the best of our knowledge, moving diffusive source tracking using particle filtering approach has not been attempted before. The posterior Cramér-Rao lower bound (PCRLB) for the moving source state estimates is also derived as a theoretical performance bound.

- We propose Dirichlet process mixture model (DPMM) [59] classifier based algorithms to estimate the location of continuous diffusive source emitting substance in a dispersive environment for both single and multiple source cases. We analytically prove convergence of the proposed algorithms in terms of total variation norm when the number of iterations in the learning process approach infinity. For the proposed algorithms, knowledge of the family of distribution of the likelihood function is enough. The non-parametric nature of DPMM framework allows for minimal knowledge of measurement data, arbitrary noise distributions, and can estimate unknown parameters of corresponding probability distribution from data itself in order to best fit the underlying the observation model.

- We address the issue of *fair* allocation of power among sensing nodes that are subjected to power constraints. We formulate and solve the problems of sequential estimation and power allocation in WSNs as a combined problem in a cooperative game theoretic framework. The proposed framework allows addressing problems of fair resource allocation for sequential estimation at the Fusion center (FC) of a wireless sensor network as a solution of a cooperative game. We propose two novel and simple game theoretic solutions to power allocation problem for sensor nodes such that the obtained solution leads to increased lifetime of the WSN tasked with estimating desired parameters. In particular, we use the concepts of Nash bargaining [60] and the Shapley function [61] for our proposed game theoretic solutions.

## 1.5 Dissertation Outline

The rest of the dissertation is organized as follows: Chapter 2 focuses on the mathematical formulations of a special physical diffusion phenomenon, i.e., an underwater oil spill scenario, and solving corresponding diffusion equations for space-time concentration distribution of the dispersed substance. In Chapter 3, the parametric estimation of a static diffusive source location has been proposed and discussed using ML and BLUE estimators for a special case of our obtained physical dispersion model in Chapter 2. Detailed analytical proofs corresponding to the consistency and asymptotic normality for the obtained ML estimator are presented when the number of sensor nodes and time samples go to infinity. The Cramér-Rao lower bound (CRLB) was also obtained as a theoretical performance bound. The proposed methods can be extended to the detection and estimation of other diffusive source estimation problem that may be encountered in homeland security applications, environmental monitoring, pollution and temperature control etc.

Parametric diffusive source localization methods presented in Chapter 3 extensively depend on the prior knowledge and sophistication of the underlying statistical model in most cases. As an alternative, in Chapter4, we explore the rich tools of nonparametric Bayesian methods, in particular Dirichlet process mixture model (DPMM) which provides us with a flexible and data-driven estimation process. We propose DPMM based static diffusive source localization algorithm and provide analytical proof of convergence. The proposed algorithm is also extended to the scenario when multiple diffusive sources of same kind are present in the diffusive field of interest.

In case of time varying parameter estimation, i.e., moving diffusive source localization, the tracking problem becomes challenging as the space-time concentration distribution at any time is affected by all past values of source position. To that end, Chapter 5 proposes a particle filter (PF) based target tracking method for moving diffusive source. To the best of our knowledge, moving diffusive source tracking using sequential Monte Carlo based PF

approach has not been attempted before. The Posterior Cramér-Rao Lower Bound (PCRLB) for the moving source state estimates is also derived as a theoretical performance bound [62].

In Chapter 6 and 7, the problem of *fair* resource allocation and lifetime maximization of WSNs, tasked with estimating desired parameters, is formulated in a cooperative game theoretic framework using the Shapley value [61] and Nash bargaining [57] concepts respectively. The proposed frameworks allow addressing problems of fair resource, i.e., power allocation for sequential estimation at the Fusion center (FC) of WSNs as solution of a cooperative games. We proposed the Shapley value and Nash bargaining based simple game theoretic solutions to power allocation problems for sensor nodes such that the obtained solution leads to increased lifetime of the WSN.

In Chapter 8, we conclude the dissertation by summarizing our works and discussing future research directions. Finally, Chapter 9 discusses some other research work done in cognitive radio area.

# Chapter 2

# Physical Modeling of Diffusion Phenomenon

The word "diffusion" means spreading, propagation or dispersion. Oxford dictionary defines it as:"*the spreading of something more widely*". It has been used in different contexts while maintaining the same literal meaning. Although diffusion phenomena have been experienced since its early days, e.g., smelling remote things, mixing of solids into liquids etc., the first related scientific record was accounted in 1827 by Scottish botanist Robert Brown [63]. Systematic experimental study of diffusion was performed by Thomas Graham in 1831 [64]. However, in 1855, Adolf Fick proposed the first mathematical laws of diffusion [65]. In this dissertation, since we are dealing with physical diffusion phenomena, we use the terms "diffusive" and/or "dispersive" to represent sources emitting diffusing substance into the environment. In physics and chemistry, diffusion is the process of matter being transported from one part of a system to another as a result of random molecular motions [65, 66]. Diffusion equation describes the dispersion of particles from a region of high concentration to regions of lower concentration due to random molecular motion.

Diffusion happens all around us, and affects our everyday lives both in positive and

negative ways, i.e., cologne or perfume permeating the air in a room, exchange of oxygen into muscles from the blood cells in the blood stream, and air pollution, water pollution etc. Among the disadvantages of physical diffusion process, the release of liquid petroleum hydrocarbon into the ocean or coastal water due to human activity has attracted tremendous attention because of its environmental, biological and economical impact. Environmental accidents, such as, BP oil disaster can severely damage the marine and wildlife habitats as well as the Gulf's fishing and tourism industries. Research in modeling and predicting the extent of such oil spill can assist in planning and emergency decision making, thereby reducing the threats and hazardous effects on environment as well as the economic cost. Over the last decades, researchers and scientists in the signal processing community have given considerable amount of attention in addressing the difficult task of diffusive source estimation, in particular, diffusive source localization [3–6]. Considering the fact that this is a diffusive source estimation and tracking problem, such research can in general be applicable in many other similar contexts such as homeland security, environmental and industrial monitoring, pollution control, servers and data center temperature monitoring as well [3, 6–12].

In this chapter, we first determine the space-time concentration distribution of the dispersion from physical modeling and mathematical formulations of a special diffusion phenomenon, i.e., an underwater oil spill scenario, considering the effect of laminar water velocity as an external force. The transport model of a substance from a diffusive source can be obtained by solving the corresponding diffusion equations.

The remainder of this chapter is organized as follows: Section 2.1 discusses problem formulation and modeling of an underwater oil spill scenario. Section 2.2 shows the validity of our obtained model through numerical simulations. The chapter concludes with summary on the obtained model and performance in section 2.3.

## 2.1 Problem Formulation and Modeling

Let us denote the concentration of the diffused substance at a position $\mathbf{r} = [x, y, z]^T$ and at time $t$ as $c(\mathbf{r}, t)$. Ignoring the effects of external forces for a source-free volume and for space-invariant diffusivity constant $\kappa$, the concentration of a dispersed substance follows the following diffusion equation [65]:

$$\frac{\partial c(\mathbf{r},t)}{\partial t} = \kappa \nabla^2 c(\mathbf{r},t) = \kappa \left( \frac{\partial^2 c(\mathbf{r},t)}{\partial x^2} + \frac{\partial^2 c(\mathbf{r},t)}{\partial y^2} + \frac{\partial^2 c(\mathbf{r},t)}{\partial z^2} \right). \tag{2.1}$$

To solve the above differential equation, appropriate boundary and initial conditions are required. We first compute the concentration for a stationary impulse point source of unit mass to obtain the Green's function. The obtained result is then extended for a continuous source by integrating the source-release rate with the Green's function. Denoting the Green's function of the impulse source as $c_G(\mathbf{r}, t)$, the concentration of a continuous point source with mass release rate $\mu(t)$ and initial release time $t_I$, can then be given by the following integral:

$$c(\mathbf{r},t) = \int_{t_I}^{t} \mu(\tau) c_G(\mathbf{r}, t-\tau) d\tau. \tag{2.2}$$

For parametric estimation case, it is to be noted that from the concentration measurements taken by the sensors, we can first estimate the source parameters of interest, and then predict its cloud evolution in space and time by inserting the estimated parameters into (2.2).

Although, the main focus of this dissertation are diffusive source localization and tracking, we introduce a special diffusion phenomenon, i.e., an underwater oil spill, to demonstrate how to model and solve for a practical diffusion phenomenon, and also to motivate the practical importance of the problem we are discussing. As shown in Fig. 2.1, we may model an under-water oil spill as diffusion occurring in a two-layer semi-infinite medium (i.e. water and air). We assume that the oil spilling source is located at the bottom (i.e. river/sea bed) at location $\mathbf{r}_0 = [x_0, y_0, z_0]^T$. The depth of water level is $0 \leq z \leq L$ with diffusivity $\kappa_w$ and

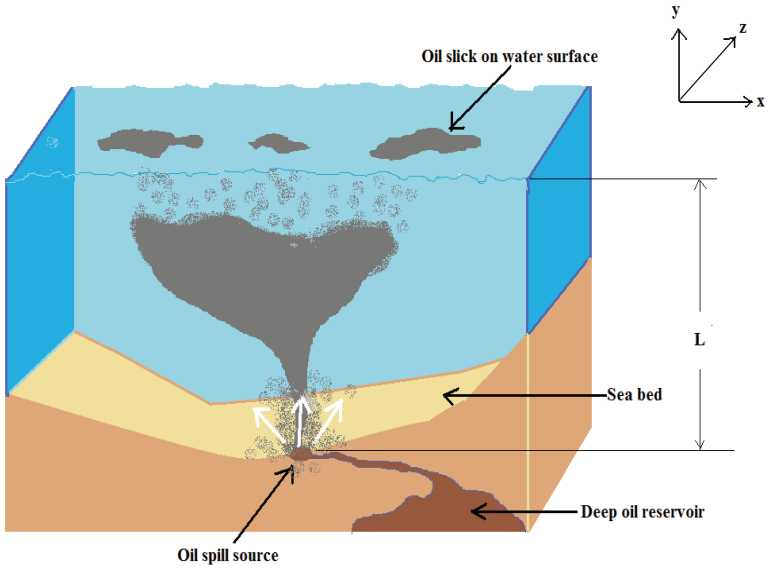



Figure 2.1: An under-water oil spill

concentration $c_w$. The same quantities for air ($z > L$) are denoted as $\kappa_a$ and $c_a$ respectively. Along the $z$-axis, we need to solve the following differential equations:

$$\frac{\partial c_w}{\partial t} = \kappa_w \frac{\partial^2 c_w}{\partial z^2}, \quad \text{for } 0 < z < L,$$

$$\frac{\partial c_a}{\partial t} = \kappa_a \frac{\partial^2 c_a}{\partial z^2}, \quad \text{for } z > L.$$

Considering only point impulse source located at $z = z_0$, where $0 \leq z_0 \leq L$ and impermeable boundary at $z = 0$, we have the following initial condition:

$$c_w(z, t) = \delta(z - z_0), \quad \text{at } t = t_I$$

and boundary conditions:

$$c_w = c_a, \quad \text{at } z = L, \tag{2.3}$$

$$\kappa_w \frac{\partial c_w}{\partial z} = \kappa_a \frac{\partial c_a}{\partial z}, \quad \text{at } z = L, \tag{2.4}$$

$$\frac{\partial c_w}{\partial z} = 0, \quad \text{at } z = 0. \tag{2.5}$$

Boundary condition (2.3) implies the continuity of concentration at the interface $z = L$. Condition (2.4) represents the fact that there is no accumulation of diffusing substance at $z = L$. Finally, the third boundary condition in (2.5) reflects the assumption that the medium at $z = 0$ is impermeable. Applying the concept of Laplace transform on the above system of partial differential equations, we can obtain the solution to the spatio-temporal concentration distribution (omitting the details in [5, 65, 66]):

$$\begin{aligned} c_w(z,t) = & \frac{1}{2\sqrt{\pi\kappa_w(t-t_I)}} \sum_{n=0}^{\infty} \rho^n \left[ \exp\left\{ -\frac{(z - z_0 - 2nL)^2}{4\kappa_w(t-t_I)} \right\} + \exp\left\{ -\frac{(z + z_0 + 2nL)^2}{4\kappa_w(t-t_I)} \right\} \right] \\ & + \frac{1}{2\sqrt{\pi\kappa_w(t-t_I)}} \sum_{n=0}^{\infty} \rho^{(n+1)} \left[ \exp\left\{ -\frac{(z - z_0 - 2(n+1)L)^2}{4\kappa_w(t-t_I)} \right\} \right. \\ & \left. + \exp\left\{ -\frac{(z + z_0 + 2(n+1)L)^2}{4\kappa_w(t-t_I)} \right\} \right], \end{aligned}$$

where $\rho = \frac{\sqrt{\kappa_w} - \sqrt{\kappa_a}}{\sqrt{\kappa_w} + \sqrt{\kappa_a}}$. As can be seen, the concentration curve can be considered to be the superimposed curve resulting from each successive reflection (from the surface layer) being superimposed on the original curve. In practice, if $\kappa_w \gg \kappa_a$, then $\rho \to 1$. Therefore we have,

$$\begin{aligned} c_w(z,t) = & \frac{1}{2\sqrt{\pi\kappa_w(t-t_I)}} \left[ \exp\left\{ -\frac{(z - z_0)^2}{4\kappa_w(t-t_I)} \right\} + \exp\left\{ -\frac{(z + z_0)^2}{4\kappa_w(t-t_I)} \right\} \right] \\ & + \frac{1}{\sqrt{\pi\kappa_w(t-t_I)}} \sum_{n=1}^{\infty} \left[ \exp\left\{ -\frac{(z - z_0 - 2nL)^2}{4\kappa_w(t-t_I)} \right\} + \exp\left\{ -\frac{(z + z_0 + 2nL)^2}{4\kappa_w(t-t_I)} \right\} \right]. \end{aligned} \tag{2.6}$$

Considering the laminar water velocity working along the X-Y plane as an external force, we have $\mathbf{v} = [v_x, v_y, 0]^T$. The diffusion equations along the x and y axes will include additional

*advection* term [65]:

$$\frac{\partial c_w}{\partial t} = \kappa_w \frac{\partial^2 c_w}{\partial x^2} - v_x \frac{\partial c_w}{\partial x}, \tag{2.7}$$

and

$$\frac{\partial c_w}{\partial t} = \kappa_w \frac{\partial^2 c_w}{\partial y^2} - v_y \frac{\partial c_w}{\partial y}. \tag{2.8}$$

Taking the Fourier transform of (2.7), we obtain:

$$\frac{dC}{dt} = \left(-\kappa_w \omega_x^2 - j v_x \omega_x\right) C, \tag{2.9}$$

where $C(\omega_x, t)$ is the Fourier transform of $c_w(x,t)$ with respect to spatial coordinate $x$. Integrating with respect to time index $t$ and taking the inverse Fourier transform, we again obtain the following:

$$\begin{aligned} c_w(x,t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\left[-\kappa_w(t-t_I)\omega_x^2 + j\{x - v_x(t-t_I) - x_0\}\omega_x\right] d\omega_x \\ &= \frac{1}{2\pi\sqrt{\kappa_w(t-t_I)}} \exp\left[-\frac{\{x - x_0 - v_x(t-t_I)\}^2}{4\kappa_w(t-t_I)}\right] \cdot \int_{-\infty}^{\infty} \exp\left[-\left(m - \frac{ja}{2}\right)^2\right] dm, \\ \therefore c_w(x,t) &= \frac{1}{2\sqrt{\pi\kappa_w(t-t_I)}} \exp\left[-\frac{\{x - x_0 - v_x(t-t_I)\}^2}{4\kappa_w(t-t_I)}\right], \end{aligned} \tag{2.10}$$

where $a = \frac{x - v_x(t-t_I) - x_0}{\sqrt{\kappa_w(t-t_I)}}$. Similarly, we can obtain the following for the $y$-axis:

$$c_w(y,t) = \frac{1}{2\sqrt{\pi\kappa_w(t-t_I)}} \exp\left[-\frac{\{y - y_0 - v_y(t-t_I)\}^2}{4\kappa_w(t-t_I)}\right]. \tag{2.11}$$

Based on our assumptions on initial and boundary conditions, and for rectangular parallelepiped space, the Green's function solution for 3-spatial-variable case is the product of the solutions of the three single spatial-variable cases with stationary impulse point source [5,65].

Therefore, the Green's function $c_G(\mathbf{r},t)$ for the space-time concentration distribution can be obtained as the product of the solutions in (2.6), (2.10) and (2.11):

$$\begin{aligned} c_G(\mathbf{r},t) &= \frac{1}{8\{\pi\kappa_w(t-t_I)\}^{3/2}}\left[\exp\left\{-\frac{|\mathbf{r}-\mathbf{v}(t-t_I)-\mathbf{r}_0|^2}{4\kappa_w(t-t_I)}\right\}+\exp\left\{-\frac{|\mathbf{r}-\mathbf{v}(t-t_I)-\mathbf{r}'|^2}{4\kappa_w(t-t_I)}\right\}\right] \\ &+\frac{1}{4\{\pi\kappa_w(t-t_I)\}^{3/2}}\exp\left[-\frac{\{x-x_0-v_x(t-t_I)\}^2}{4\kappa_w(t-t_I)}-\frac{\{y-y_0-v_y(t-t_I)\}^2}{4\kappa_w(t-t_I)}\right] \\ &\times\sum_{n=1}^{\infty}\left[\exp\left\{-\frac{\{z-z_0-2nL\}^2}{4\kappa_w(t-t_I)}\right\}+\exp\left\{-\frac{\{z+z_0+2nL\}^2}{4\kappa_w(t-t_I)}\right\}\right], \end{aligned} \tag{2.12}$$

where $\mathbf{r}_0=[x_0,y_0,z_0]^T$ and $\mathbf{r}'=[x_0,y_0,-z_0]^T$. Considering the source mass release rate to be constant $\mu(t)=\mu$, the final solution for concentration of oil diffusion in water for stationary continuous source with mass rate of $\mu(t)$ can be obtained from (2.2):

$$c(\mathbf{r},t) = \mu\int_{t_I}^{t} c_G(\mathbf{r},t-\tau)d\tau = c_1(\mathbf{r},t)+c_2(\mathbf{r},t)+c_3(\mathbf{r},t)+c_4(\mathbf{r},t), \tag{2.13}$$

where

$$\begin{aligned} c_1(\mathbf{r},t) &= \frac{\mu}{8\pi\kappa_w|\mathbf{r}-\mathbf{r}_0|}\exp\left\{\frac{(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{v}}{2\kappa_w}\right\}\left[\exp\left\{\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\operatorname{erfc}\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{2\sqrt{\kappa_w(t-t_I)}}+|\mathbf{v}|\sqrt{\frac{t-t_I}{4\kappa_w}}\right\}\right. \\ &\left.+\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\times\operatorname{erfc}\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{2\sqrt{\kappa_w(t-t_I)}}-|\mathbf{v}|\sqrt{\frac{t-t_I}{4\kappa_w}}\right\}\right], \end{aligned}$$

$$\begin{aligned} c_2(\mathbf{r},t) &= \frac{\mu}{8\pi\kappa_w|\mathbf{r}-\mathbf{r}'|}\exp\left\{\frac{(\mathbf{r}-\mathbf{r}')\cdot\mathbf{v}}{2\kappa_w}\right\}\left[\exp\left\{\frac{|\mathbf{r}-\mathbf{r}'||\mathbf{v}|}{2\kappa_w}\right\}\operatorname{erfc}\left\{\frac{|\mathbf{r}-\mathbf{r}'|}{2\sqrt{\kappa_w(t-t_I)}}+|\mathbf{v}|\sqrt{\frac{t-t_I}{4\kappa_w}}\right\}\right. \\ &\left.+\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}'||\mathbf{v}|}{2\kappa_w}\right\}\times\operatorname{erfc}\left\{\frac{|\mathbf{r}-\mathbf{r}'|}{2\sqrt{\kappa_w(t-t_I)}}-|\mathbf{v}|\sqrt{\frac{t-t_I}{4\kappa_w}}\right\}\right], \end{aligned}$$

$$c_3(\mathbf{r},t) = \frac{\mu}{4(\pi\kappa_w)^{3/2}}\sum_{n=1}^{\infty}\int_0^{t-t_I}\tau^{-3/2}\exp\left\{-\frac{(x-x_0-v_x\tau)^2+(y-y_0-v_y\tau)^2+(z-z_0-2nL)^2}{4\kappa_w\tau}\right\}d\tau,$$

$$c_4(\mathbf{r},t) = \frac{\mu}{4(\pi\kappa_w)^{3/2}}\sum_{n=1}^{\infty}\int_0^{t-t_I}\tau^{-3/2}\exp\left\{-\frac{(x-x_0-v_x\tau)^2+(y-y_0-v_y\tau)^2+(z+z_0+2nL)^2}{4\kappa_w\tau}\right\}d\tau.$$

Derivation to (2.13) is given in Appendix 2.3. For the sake of simplicity, from here on, we denote the diffusivity constant $\kappa_w = \kappa$.

### 2.1.1 Moving Diffusive Source

For a moving diffusive source emitting substance continuously in a semi-infinite medium similar to our case, space-time concentration distribution can be obtained using the concept of convolution integral from the Green's function solution corresponding to stationary impulsive source. In this case, substance concentration at any time instant is affected by all the past values of source position and release rate. Therefore, time-cumulation effect on the concentrations has to be considered to obtain complete physical model. For a moving diffusive source continuously releasing substance at a mass rate $\mu(t)$, the space-time concentration distribution in a semi-infinite medium can be obtained for a given Green's function $c_G(\mathbf{r},t)$ using the following integral:

$$c(\mathbf{r},t) = \int_{t_I}^{t}\mu(\tau)c_G\left(\mathbf{r}-\mathbf{r}_0(\tau),t-\tau\right)d\tau, \tag{2.14}$$

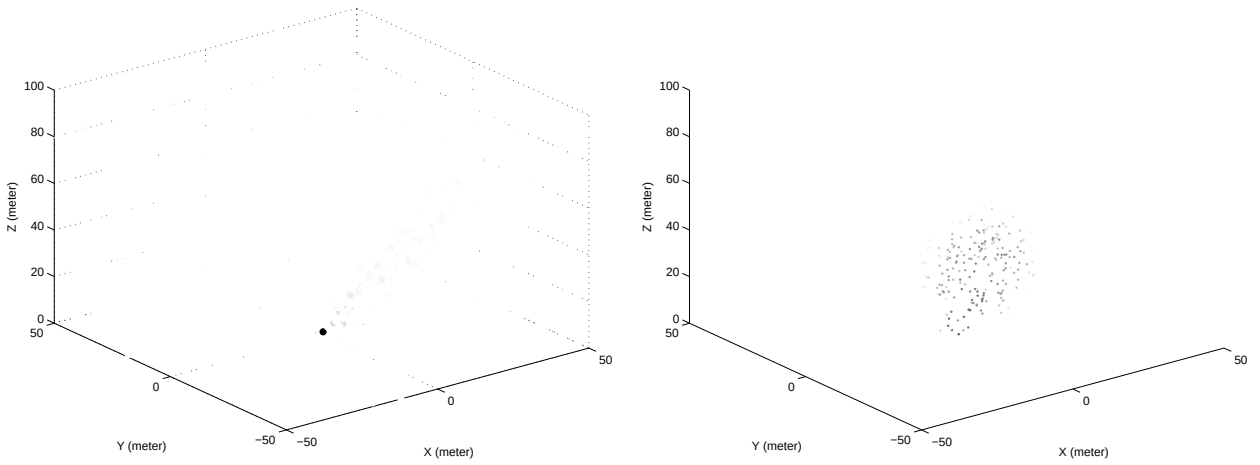
where $\mathbf{r}_0(t) = [x_0(t), y_0(t), z_0(t)]^T$ represents the source moving path. The advantage of solving the physical diffusion model corresponding to a moving diffusive source using (2.14) is that the initial, boundary, and other necessary conditions can be taken into account to solve for the stationary case in the first step before extending it to the moving source case.

## 2.2 Simulation Results

In the following, we give simple examples of how concentration of chemical substance dispersion evolve over time and/or space. In these examples, we show the concentration dis-

tribution of a point source at a location $\mathbf{r}_0$ in a homogeneous semi-infinite medium with impermeable boundary, releasing diffusing chemical substance with $\mu$ Kg/s and medium diffusivity $\kappa$ $m^2/s$ at $t_I$ sec. The parameters used for simulation in this section are: mass release rate $\mu = 10^3 Kg/s$, diffusivity constant $\kappa = 25m^2/s$, initial release time $t_I = 0$ sec and velocity vector $\mathbf{v} = [50, 50, 0]m/s$. The source is assumed to be located at $\mathbf{r}_0 = [0, 0, 0]^T$ and the depth of water is taken to be $L = 100$m from the sea bed.

Figure 2.2 shows the spatial concentration distribution for two different time instants $t = 1$ and $t = 100$ sec. It can be seen from Fig. 2.2 that as the source emitting the chemical substance is located at the origin, concentration is high near the origin at $t = 1$ sec (magnitude of concentration is proportional to darkness). It is interesting to see that since laminar water flow is assumed to be only active in the x and y directions, concentration increases more along the positive x-y plane with the increase in time.




(a) (b)

Figure 2.2: Concentration distribution in space ($x-y-z$ coordinates) at times (a) $t = 1$ sec, and (b) $t = 100$ sec with velocity vector $\mathbf{v} = [50, 50, 0]$ m/s. (Magnitude of concentration is proportional to darkness)

## 2.3 Conclusion

In this chapter, we introduced a special diffusion phenomenon, i.e., an underwater oil spill, and obtained spatial and temporal distribution of the substance concentration by solving corresponding diffusion equations with appropriate initial, boundary conditions, and pragmatic assumptions. Though oil spill problem has been discussed as a special case, in general, the solution techniques used in this chapter, are applicable (with appropriate modifications) to many other similar contexts as well, including environmental and industrial monitoring, pollution control, biochemical dispersion, temperature monitoring etc. Differences in modeling and solution methods occur during the mathematical formulations of the physical diffusion phenomena leading to different initial and boundary conditions. While solving for the spatio-temporal concentration distribution corresponding to an underwater oil spill scenario, our main objective was to obtain an analytical solution to the diffusion problem, rather than using non-model based sophisticated numerical techniques. Simulation results are provided to demonstrate the effectiveness of the obtained model.

# Appendix 2A

### Derivation of Spatio-Temporal Concentration Distribution in (2.13)

To derive and verify the spatio-temporal concentration distribution in (2.13), the Green's function $c_G(\mathbf{r},t)$ in (2.12) can be written as $c_G(\mathbf{r},t) = c_1'(\mathbf{r},t) + c_2'(\mathbf{r},t) + c_3'(\mathbf{r},t) + c_4'(\mathbf{r},t)$, where

$$\begin{aligned} c_1'(\mathbf{r},t) &= \frac{1}{8\{\pi\kappa_w(t-t_I)\}^{3/2}} \exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0-\mathbf{v}(t-t_I)|^2}{4\kappa_w(t-t_I)}\right\}, \\ c_2'(\mathbf{r},t) &= \frac{1}{8\{\pi\kappa_w(t-t_I)\}^{3/2}} \exp\left\{-\frac{|\mathbf{r}-\mathbf{r}'-\mathbf{v}(t-t_I)|^2}{4\kappa_w(t-t_I)}\right\}, \end{aligned}$$

$$
\begin{aligned}
c_3'(\mathbf{r},t) &= \frac{1}{4\{\pi\kappa_w(t-t_I)\}^{3/2}}\exp\left[-\frac{\{x-x_0-v_x(t-t_I)\}^2}{4\kappa_w(t-t_I)}-\frac{\{y-y_0-v_y(t-t_I)\}^2}{4\kappa_w(t-t_I)}\right] \\
&\quad\times\sum_{n=1}^{\infty}\exp\left\{-\frac{(z-z_0-2nL)^2}{4\kappa_w(t-t_I)}\right\},\ \text{and} \\
c_4'(\mathbf{r},t) &= \frac{1}{4\{\pi\kappa_w(t-t_I)\}^{3/2}}\exp\left[-\frac{\{x-x_0-v_x(t-t_I)\}^2}{4\kappa_w(t-t_I)}-\frac{\{y-y_0-v_y(t-t_I)\}^2}{4\kappa_w(t-t_I)}\right] \\
&\quad\times\sum_{n=1}^{\infty}\exp\left\{-\frac{(z+z_0+2nL)^2}{4\kappa_w(t-t_I)}\right\}.
\end{aligned}
$$

Therefore, performing change of variables, we can rewrite $c_1(\mathbf{r},t)$ in (2.13) as follows:

$$
\begin{aligned}
c_1(\mathbf{r},t) &= \mu\int_{t_I}^{t} c_1'(\mathbf{r},t-\tau)d\tau, \\
&= \int_{t_I}^{t}\frac{\mu}{8\{\pi\kappa_w(t-\tau+t_I)\}^{3/2}}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0-\mathbf{v}(t-\tau+t_I)|^2}{4\kappa_w(t-\tau+t_I)}\right\}d\tau, \\
&= \int_{0}^{t-t_I}\frac{\mu}{(4\pi\kappa_w\tau)^{3/2}}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0-\mathbf{v}\tau|^2}{4\kappa_w\tau}\right\}d\tau, \qquad (2.15)
\end{aligned}
$$

To prove that (2.15) indeed translates into the expression given in (2.13), we will use the concept of first fundamental theorem of calculus [67]. Since $c_1'(\mathbf{r},t)$ is a continuous real-valued function within the limits of the integral, derivate of the expression given in (2.13) will be taken to verify (2.15). Replacing $\gamma = t - t_I$ and assuming $F(\mathbf{r},t) = c_1(\mathbf{r},t)$ in (2.13), we have:

$$
\begin{aligned}
F(\mathbf{r},\gamma) &= \frac{\mu}{8\pi\kappa_w|\mathbf{r}-\mathbf{r}_0|}\exp\left\{\frac{(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{v}}{2\kappa_w}\right\}\left[\exp\left\{\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\operatorname{erfc}\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{2\sqrt{\kappa_w\gamma}}+|\mathbf{v}|\sqrt{\frac{\gamma}{4\kappa_w}}\right\}\right. \\
&\quad\left.+\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\times\operatorname{erfc}\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{2\sqrt{\kappa_w\gamma}}-|\mathbf{v}|\sqrt{\frac{\gamma}{4\kappa_w}}\right\}\right].
\end{aligned}
$$

Since $\frac{d}{dz}\text{erfc}(z) = -\frac{2}{\sqrt{\pi}}\exp(-z^2)$, we can obtain the following:

$$\begin{aligned}
\frac{\partial F(\mathbf{r},\gamma)}{\partial \gamma} &= \frac{\mu}{4\pi^{3/2}\kappa_w|\mathbf{r}-\mathbf{r}_0|}\exp\left\{\frac{(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{v}}{2\kappa_w}\right\}\left[\exp\left\{\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{4\sqrt{\kappa_w}\gamma^{3/2}}-\frac{|\mathbf{v}|}{4\sqrt{\gamma\kappa_w}}\right\}\right.\\
&\quad\times\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa_w\gamma}-\frac{|\mathbf{v}|^2\gamma}{4\kappa_w}-\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}+\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\\
&\quad\left.\times\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{4\sqrt{\kappa_w}\gamma^{3/2}}+\frac{|\mathbf{v}|}{4\sqrt{\gamma\kappa_w}}\right\}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa_w\gamma}-\frac{|\mathbf{v}|^2\gamma}{4\kappa_w}+\frac{|\mathbf{r}-\mathbf{r}_0||\mathbf{v}|}{2\kappa_w}\right\}\right],\\
&= \frac{\mu}{4\pi^{3/2}\kappa_w|\mathbf{r}-\mathbf{r}_0|}\exp\left\{\frac{(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{v}}{2\kappa_w}\right\}\left[\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{4\sqrt{\kappa_w}\gamma^{3/2}}-\frac{|\mathbf{v}|}{4\sqrt{\gamma\kappa_w}}\right\}\right.\\
&\quad\times\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa_w\gamma}-\frac{|\mathbf{v}|^2\gamma}{4\kappa_w}\right\}+\left\{\frac{|\mathbf{r}-\mathbf{r}_0|}{4\sqrt{\kappa_w}\gamma^{3/2}}+\frac{|\mathbf{v}|}{4\sqrt{\gamma\kappa_w}}\right\}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa_w\gamma}-\frac{|\mathbf{v}|^2\gamma}{4\kappa_w}\right\}\left.\right],\\
&= \frac{\mu}{(4\pi\kappa_w\gamma)^{3/2}}\exp\left\{\frac{(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{v}}{2\kappa_w}\right\}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa_w\gamma}-\frac{|\mathbf{v}|^2\gamma}{4\kappa_w}\right\},\\
&= \frac{\mu}{(4\pi\kappa_w\gamma)^{3/2}}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2-2\gamma(\mathbf{r}-\mathbf{r}_0)\cdot\mathbf{v}+|\mathbf{v}|^2\gamma^2}{4\kappa_w\gamma}\right\},\\
\therefore \frac{\partial F(\mathbf{r},\gamma)}{\partial \gamma} &= \frac{\mu}{(4\pi\kappa_w\gamma)^{3/2}}\exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0-\mathbf{v}\gamma|^2}{4\kappa_w\gamma}\right\}. \qquad (2.16)
\end{aligned}$$

Hence, the resulting expression for $c_1(\mathbf{r},t)$ in (2.13) is valid. The expression for $c_3(\mathbf{r},t)$ can be obtained as follows:

$$\begin{aligned}
c_3(\mathbf{r},t) &= \mu\int_{t_I}^{t} c_3'(\mathbf{r},t-\tau)d\tau,\\
&= \int_{t_I}^{t}\frac{\mu}{4\{\pi\kappa_w(t-\tau+t_I)\}^{3/2}}\exp\left[-\frac{\{x-x_0-v_x(t-\tau+t_I)\}^2}{4\kappa_w(t-\tau+t_I)}-\frac{\{y-y_0-v_y(t-\tau+t_I)\}^2}{4\kappa_w(t-\tau+t_I)}\right]\\
&\quad\times\sum_{n=1}^{\infty}\exp\left\{-\frac{(z-z_0-2nL)^2}{4\kappa_w(t-\tau+t_I)}\right\}d\tau,\\
&= \frac{\mu}{4(\pi\kappa_w)^{3/2}}\sum_{n=1}^{\infty}\int_0^{t-t_I}\tau^{-3/2}\exp\left\{-\frac{(x-x_0-v_x\tau)^2+(y-y_0-v_y\tau)^2+(z-z_0-2nL)^2}{4\kappa_w\tau}\right\}d\tau,
\end{aligned}$$

Similarly, we can also verify the expressions for $c_2(\mathbf{r},t)$ and $c_4(\mathbf{r},t)$. Therefore, the spatio-temporal concentration distribution $c(\mathbf{r},t)$ given in (2.13) is valid. ■

# Chapter 3

# Parametric Diffusive Source Localization

## 3.1 Introduction

Source localization/tracking using WSN is an area of active research interest in the signal processing community [28, 29]. Interesting research on diffusive source localization problems can be found in [3–6, 30–32], where biochemical concentration distribution in space and time for different types of diffusive sources and models are estimated in various environments. The pros and cons of the existing literature on diffusive source localization have been discussed in detail in section 1.3 of Chapter 1. In this chapter, we propose and discuss two parametric estimation based static diffusive source localization methods using the ML and BLUE estimators. We provide detailed analytical proofs corresponding to the consistency and asymptotic normality for the obtained ML estimator are presented when the number of sensor nodes and time samples go to infinity. The Cramér-Rao lower bound (CRLB) [40] was also obtained as a theoretical performance bound.

The chapter is organized as follows: Section 3.2 and 3.3 discuss, respectively, modeling of

an underwater oil spill scenario and measurement model for static diffusive source localization using sensor network. The proposed statistical methods for static diffusive source localization are discussed in section 3.4 and 3.5. Section 3.6 shows the validity and effectiveness of our proposed methods for static diffusive source localization by comparing numerical simulations with analytically obtained theoretical bound . Finally, section 3.7 concludes the chapter by summarizing our results.

## 3.2 Physical Model

For simplicity of exposition, we consider a special case of our obtained physical model in Chapter 2, when an oil spill occurs in an infinite ($L \to \infty$) underwater medium with no external force. Using the concept of Fourier and inverse-Fourier transform, and extending (2.10) we can solve for the following Green's function for the aforementioned case:

$$\begin{aligned} c_G(\mathbf{r},t) &= \left[\frac{\exp\left\{-\frac{(x-x_0)^2}{4\kappa(t-t_I)}\right\}}{2\sqrt{\pi\kappa(t-t_I)}}\right] \times \left[\frac{\exp\left\{-\frac{(y-y_0)^2}{4\kappa(t-t_I)}\right\}}{2\sqrt{\pi\kappa(t-t_I)}}\right] \times \left[\frac{\exp\left\{-\frac{(z-z_0)^2}{4\kappa(t-t_I)}\right\}}{2\sqrt{\pi\kappa(t-t_I)}}\right], \\ &= \frac{1}{8\left\{\pi\kappa(t-t_I)\right\}^{\frac{3}{2}}} \exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa(t-t_I)}\right\}. \end{aligned} \tag{3.1}$$

Considering the source mass release rate to be constant $\mu(t) = \mu$, the final solution for concentration of oil diffusion in water for stationary continuous source with mass rate of $\mu(t)$ can be obtained from (2.2) using (3.1), [65]:

$$\begin{aligned} c(\mathbf{r},t) &= \mu \int_{t_I}^{t} c_G(\mathbf{r}, t-\tau)d\tau = \int_0^{t-t_I} \frac{\mu}{(4\pi\kappa\tau)^{\frac{3}{2}}} \exp\left\{-\frac{|\mathbf{r}-\mathbf{r}_0|^2}{4\kappa\tau}\right\} d\tau, \\ &= \frac{\mu}{4\pi\kappa|\mathbf{r}-\mathbf{r}_0|}\text{erfc}\left(\frac{|\mathbf{r}-\mathbf{r}_0|}{2\sqrt{\kappa(t-t_I)}}\right). \end{aligned} \tag{3.2}$$

In this case, the concentration at any position $\mathbf{r}_j$ at time $t_k$ is reduced to the following expression [3, 65]:

$$c(\mathbf{r}_j, t_k) = \frac{\mu}{4\pi\kappa|\mathbf{r}_j - \mathbf{r}_0|}\text{erfc}\left(\frac{|\mathbf{r}_j - \mathbf{r}_0|}{2\sqrt{\kappa(t_k - t_I)}}\right), \tag{3.3}$$

where erfc(.) is the complementary error function. Assuming that the spatio-temporal concentration distribution given in (3.3) represents the underlying diffusion phenomenon, we transform the transport model into a statistical measurement model in order to facilitate the use of the proposed methods.

## 3.3 Measurement Model

We consider a WSN consisting of a fusion center (FC) and $N$ spatially distributed biochemical static sensor nodes capable of sensing in dispersive environment. For practical consideration, we assume that the $N$ distributed sensors are located in a rectangular volume in space such that $\mathbf{r}_j = [x_j, y_j, z_j]^T \in \Lambda, \forall j \in \{1, 2, ..., N\}$, where $\Lambda = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2] \subseteq \mathbb{R}^3$. It is also assumed that the source-to-sensor distances are much higher than the source and sensor dimensions. Each sensor node takes measurements at times $t_k; \forall k \in \{1, 2, ..., T\}$, where $T$ is the total number of time samples. We may obtain a measurement model for a sensor at a position $\mathbf{r}_j$ and at time $t_k$ as:

$$y(\mathbf{r}_j, t_k) = c(\mathbf{r}_j, t_k) + e(\mathbf{r}_j, t_k) + b, \tag{3.4}$$

where $c(\mathbf{r}_j, t_k)$ is the concentration of interest, $b$ is a bias term, and $e(\mathbf{r}_j, t_k) \sim \mathcal{N}(0, \sigma^2)$ is the sensor noise assumed to be independent in both time and space. For the sake of brevity, (3.4) can be rewritten in the simplified form as:

$$y_{j,k} = c_{j,k}(\boldsymbol{\theta}) + e_{j,k} + b, \tag{3.5}$$

where $y_{j,k} = y(\mathbf{r}_j, t_k)$, $e_{j,k} = e(\mathbf{r}_j, t_k)$, $c_{j,k}(\boldsymbol{\theta}) = c(\mathbf{r}_j, t_k)$, $\boldsymbol{\theta} \in \mathbb{R}^{n\times 1}$ is the unknown source and medium parameter vector that we are interested to estimate, and $b$ is the bias or clutter term representing the sensor's response to foreign substances that may be present in a diffusive field of interest. The bias term is assumed to be space and time-invariant, such that the foreign substances interfering with the actual measurements are in steady state. If we want to localize a static diffusive source, then only $[x_0, y_0, z_0]$ are the parameters of interest. It is to be noted that some of the parameters, such as, the diffusivity constant $\kappa$, bias term $b$ and noise variance $\sigma^2$ can be measured at the calibration stage, thereby reducing the cost of computation during the detection/estimation phase.

We assume that the sensor nodes are in sleep mode until they are activated by some central control (i.e. FC) due to a possible release of substance from a diffusive source. The activated sensor nodes take measurements of substance concentration at time instants $t_k$'s and then return to sleep mode. For $N$ number of nodes in a WSN and each node taking $T$ number of time samples of the substance concentrations at their respective locations, let $\boldsymbol{y} \in \mathbb{R}^{NT\times 1}$ be the measurement vector received at the FC.

## 3.4 Maximum-likelihood (ML) Based Parameter Estimation

In this section, we use the maximum-likelihood (ML) estimator to estimate the unknown parameter vector $\boldsymbol{\theta}$. From the measurement model discussed in section 3.3, if the bias parameter $b$ is known, conditional pdf of the measurement taken by the $j$-th node at $t_k$ time instant $p(y_{jk}|\boldsymbol{\theta}) \sim \mathcal{N}(c_{jk}(\boldsymbol{\theta}) + b, \sigma^2)$. Hence, the log-likelihood function formed at the FC

can be written as:

$$\begin{aligned}\mathcal{L} &= \sum_{j=1}^{N}\sum_{k=1}^{T}\log p(y_{jk}|\boldsymbol{\theta}) \\ &= -\frac{NT}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{j=1}^{N}\sum_{k=1}^{T}\left(y_{j,k} - c_{j,k}(\boldsymbol{\theta}) - b\right)^2. \end{aligned} \tag{3.6}$$

The log-likelihood equations are obtained by $\frac{\partial\mathcal{L}}{\partial\boldsymbol{\theta}}$:

$$\sum_{j=1}^{N}\sum_{k=1}^{T}\left(y_{j,k} - c_{j,k}(\boldsymbol{\theta}) - b\right)\left[\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial\theta_u}\right]\Bigg|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0, \tag{3.7}$$

for $u = 1, 2, 3$, where $\theta_u$ is the $u$-th element of $\boldsymbol{\theta}$, and

$$\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial\theta_u} = \frac{\mu\left[\mathbf{r}_j(u) - \mathbf{r}_0(u)\right]}{4\pi\kappa|\mathbf{r}_j - \mathbf{r}_0|^2}\left[\frac{\operatorname{erfc}\left(\frac{|\mathbf{r}_j-\mathbf{r}_0|}{2\sqrt{\kappa(t_k-t_I)}}\right)}{|\mathbf{r}_j - \mathbf{r}_0|} + \frac{\exp\left\{-\frac{|\mathbf{r}_j-\mathbf{r}_0|^2}{4\kappa(t_k-t_I)}\right\}}{\sqrt{\pi\kappa(t_k - t_I)}}\right]. \tag{3.8}$$

Since the system of equations in (3.7) is nonlinear, there is no closed-form solution to it. We can obtain an ML estimation of the source location using any suitable nonlinear optimization technique. In this case, (3.7) is solved using simplex search algorithm [68].

The Cramér-Rao lower bound (CRLB) provides a lower limit on the mean-squared estimation error of an unbiased estimator of non-random parameter [40]. CRLB in this case can be obtained as CRLB $\geq \mathbf{I}_{\boldsymbol{\theta}}^{-1}$, where $\mathbf{I}_{\boldsymbol{\theta}} \in \mathbb{R}^{3\times3}$ is the Fisher information matrix (FIM) formed at the FC. The $u$-$v$th element of the FIM can be found as:

$$\begin{aligned}\left[\mathbf{I}_{\boldsymbol{\theta}}\right]_{u,v} &= \mathbb{E}\left[\left\{\frac{\partial}{\partial\theta_u}\log p(\boldsymbol{y}|\boldsymbol{\theta})\right\}\left\{\frac{\partial}{\partial\theta_v}\log p(\boldsymbol{y}|\boldsymbol{\theta})\right\}\right], \\ &= \frac{1}{\sigma^4}\mathbb{E}\left[\left(\sum_{j=1}^{N}\sum_{k=1}^{T}\left(y_{jk} - c_{jk}(\boldsymbol{\theta}) - b\right)\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}\right)\left(\sum_{j=1}^{N}\sum_{k=1}^{T}\left(y_{jk} - c_{jk}(\boldsymbol{\theta}) - b\right)\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\}\right)\right], \\ &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial\theta_u}\right\}\left\{\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial\theta_v}\right\}, \end{aligned} \tag{3.9}$$

where (3.9) was obtained using the independence assumption of observations in space and in time.

A sequence of estimators $\hat{\boldsymbol{\theta}}_n$ to an unknown parameter vector $\boldsymbol{\theta}$ is said to be consistent if the sequence converges *in probability* to $\boldsymbol{\theta}$, i.e., $\lim_{n\to\infty} \hat{\boldsymbol{\theta}}_n = \boldsymbol{\theta}$, where $n$ is the sample size [40]. It is desirable to have a consistent MLE as consistency ensures that for large data sets, the MLE will converge to the true parameter. The obtained MLE to our source localization problem is consistent when the number of sensor nodes in any non-negligible open subset of $\Lambda = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2] \subseteq \mathbb{R}^3$ and time samples go to infinity.

*Theorem 1*: If the number of sensors $N$ increases in such a way that for any open subset of $\Lambda = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2] \subseteq \mathbb{R}^3$ having positive area, the number of sensors $N$ and/or the number of time samples $T$ tend to infinity, the obtained ML estimator is consistent.

*Proof*: See Appendix 3A.

Once consistency for the obtained MLE is established, the next important thing is to check the asymptotic normality. An asymptotically normal estimator is a consistent estimator whose distribution around the true parameter $\boldsymbol{\theta}$ approaches a normal distribution with standard deviation shrinking in proportion to $1/\sqrt{n}$ as the sample size $n$ grows, i.e., $\sqrt{nI_{\boldsymbol{\theta}}}\left(\hat{\boldsymbol{\theta}}_n - \boldsymbol{\theta}\right) \longrightarrow \mathcal{N}\left(\mathbf{0}, \mathbf{I}^{-1}\right)$, where $\mathbf{I}_{\boldsymbol{\theta}}$ and $\mathbf{I}$ are the Fisher information and identity matrices respectively [40]. It ensures that the estimator not only converges to the unknown parameter, but it converges fast enough at a rate $1/\sqrt{n}$. We address this issue with the following theorem on asymptotic normality.

*Theorem 2*: If the number of sensors $N$ and time samples $T$ increase as in Theorem 1, then for a true parameter vector $\boldsymbol{\theta}_0 \in \mathring{\Lambda}$, where $\mathring{\Lambda} \subset \Lambda$ is an open subset of $\Lambda$, the following is true

$$\sqrt{NT}\left(\hat{\boldsymbol{\theta}}_{ML}\left(\boldsymbol{y}\right) - \boldsymbol{\theta}_0\right) \longrightarrow \mathcal{N}\left(\mathbf{0}, \left(\bar{\mathbf{I}}_{\boldsymbol{\theta}_0}\right)^{-1}\right),$$

*in distribution* where the $(u, v)$-th element of the matrix $\bar{\mathbf{I}}_{\boldsymbol{\theta}}$ is given by

$$\left[\bar{\mathbf{I}}_{\boldsymbol{\theta}}\right]_{u,v} = \lim_{N,T\to\infty} \frac{1}{\sigma^2 NT} \sum_{j=1}^{N}\sum_{k=1}^{T} \left\{\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial \theta_u}\right\} \left\{\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial \theta_v}\right\}.$$

*Proof*: See Appendix 3B.

## 3.5 Best Linear Unbiased Estimator (BLUE) Based Source Localization

The advantage of using the BLUE for static diffusive source localization is that there is no constraints on the PDF, and also knowing only the mean and covariance of the measurements are enough. However, observations have to be linear for performing the BLUE algorithm. In this section, we assume that the distributed sensing nodes are capable of estimating their respective distances from the source using BLUE.

Since the complementary error function can be approximated as $\text{erfc}(z) \approx 1 - \frac{2}{\sqrt{\pi}}z$, hence our observation model for $j$-th node at the $k$-th time instant can be linearized in terms of the inverse of the source-to-node distances from (3.5) and (3.3):

$$\begin{aligned} y_{j,k} &\approx \frac{\mu|\mathbf{r}_j - \mathbf{r}_0|^{-1}}{4\pi\kappa} + \left[b - \frac{\mu}{4\sqrt{\pi^3\kappa^3(t_k - t_I)}}\right] + e_{j,k} \\ &= h d_j^{inv} + a_k + e_{j,k}, \end{aligned} \tag{3.10}$$

where $h = \frac{\mu}{4\pi\kappa}$, $d_j^{inv} = |\mathbf{r}_j - \mathbf{r}_0|^{-1}$ and $a_k = b - \frac{\mu}{4\sqrt{\pi^3\kappa^3(t_k - t_I)}}$. Since all the parameters are known except for the diffusive source location, we can write $\tilde{y}_{j,k} = y_{j,k} - a_k = h d_j^{inv} + e_{j,k}$.

Therefore, the observation vector formed at the $j$-th node can be written as:

$$\tilde{\boldsymbol{y}}_j = \begin{pmatrix} y_{i1} - a_1 \\ y_{i2} - a_2 \\ \vdots \\ y_{iT} - a_T \end{pmatrix} = \boldsymbol{h} d_j^{inv} + \boldsymbol{e}_j, \tag{3.11}$$

where $\boldsymbol{h}$ is a column vector of all $h$'s and $\boldsymbol{e}_j = [e_{i1}, e_{i2}, \ldots, e_{iT}]^T$. Since $e_{j,k} \sim \mathcal{N}(0, \sigma^2)$ for $\forall j, k$ and measurement noise is assumed to be independent and identically distributed across space and time, hence the covariance matrix of $\tilde{\boldsymbol{y}}_j$ is $\tilde{\Sigma}_j = \text{diag}(\sigma^2, \sigma^2, \ldots, \sigma^2) \in \mathbb{R}^{T \times T}$. The optimal BLUE estimator formed at $j$-th node is given by

$$\hat{d}_j^{inv} = \frac{\boldsymbol{h}^T \tilde{\Sigma}_j^{-1} \tilde{\boldsymbol{y}}_j}{\boldsymbol{h}^T \tilde{\Sigma}_j^{-1} \boldsymbol{h}}, \tag{3.12}$$

with estimator variance $V_j = \left( \boldsymbol{h}^T \tilde{\Sigma}_j^{-1} \boldsymbol{h} \right)^{-1}$.

After the distributed nodes estimate their respective distances $\hat{d}_j = |\mathbf{r}_j - \mathbf{r}_0|$ from the source using BLUE, all nodes send $\hat{d}_j$'s to the fusion center (FC) for further processing. It is to be noted that the source-to-node distance estimation can also be performed at the FC. Signal received at the FC from the $j$-th node can be expressed as $f_j = \hat{d}_j + w_j$, where $w_j$ is normally distributed with mean 0 and variance $\sigma_m^2$. For $N$ number of nodes, the data vector available at the FC can be written as:

$$\boldsymbol{F} = [f_1, f_2, \ldots, f_N]^T = \hat{\boldsymbol{D}} + \boldsymbol{w},$$

where $\hat{\boldsymbol{D}} = \left[ \hat{d}_1, \hat{d}_2, \ldots, \hat{d}_N \right]^T$ with $d_j = \sqrt{(x_j - x_0)^2 + (y_j - y_0)^2 + (z_j - z_0)^2}$, and $\boldsymbol{w} = [w_1, w_2, \ldots, w_N]^T$. The data vector $\boldsymbol{F}$ formed at the FC can be used to estimate the diffusive source location using the nonlinear least-square approach:

$$\hat{\mathbf{r}}_0 = \arg \min_{\mathbf{r}_0 = [x_0, y_0, z_0]} \left|\left| \boldsymbol{F} - \hat{\boldsymbol{D}} \right|\right|_2^2. \tag{3.13}$$

To solve for the source location from (3.13), simplex search algorithm [68] has been used.

## 3.6 Simulation Results

### 3.6.1 Static Diffusive Source Localization

Here, we show simulation results in estimating the location of a static diffusive source using the proposed MLE and BLUE based methods from the concentration observations taken by the sensing nodes. For the sake of simplicity, we consider a 2D diffusive field volume of $\Lambda = [-50, 50] \times [-50, 50]$ m$^2$. We assume that the sensors are placed in a uniform 2D grid such that the distance between adjacent sensors along the same ordinate is approximately 14.3 m. Parameters used for simulations are: number of nodes $N$ =64, $\mathbf{r}_0 = [0, 0]^T$, $\mu$= 1000 Kg/s, $b = 10^{-4}$ Kg/$m^2$, $t_I = 0$ sec and $\kappa = 25m^2/s$. The observation noise is assumed to have Gaussian distribution with mean 0 and variance $\sigma^2 = 1 \times 10^{-4}$ Kg/$m^2$. Total number of random realizations used for simulations is 100. The measurements are taken at every 0.5 sec time-step starting from 0.5 sec and ending at 30 sec. In case of BLUE estimator, received noise variance at the fusion center is assumed to be $\sigma_m^2 = 0.01, 10\ m^2$.

Figure 3.1 shows the normalized mean-squared-error (MSE) and CRLB (in dB) with the increase in the number of nodes and samples. The normalized MSE and CRLB are obtained by dividing each with the diffusive field volume. As one would except, estimation error decreases as more distributed nodes and samples are considered for estimation purpose. Since it is a 2D location estimation problem, we at least need 3 nodes to determine the source location correctly. It is interesting to note that the estimation performance is slightly better than the CRLB in some cases. This is due to the fact that the ML estimator in this case is biased (suggested from simulation), and thus it can outperform the CRLB by trading variance for bias. In this particular case, the continuous diffusive source can be localized with a resolution of less than 12 cm.

The estimated source location using the BLUE estimator is shown in Figure 3.2 as a function of the number of nodes and time samples for different values of $\sigma_m^2$'s. As one would

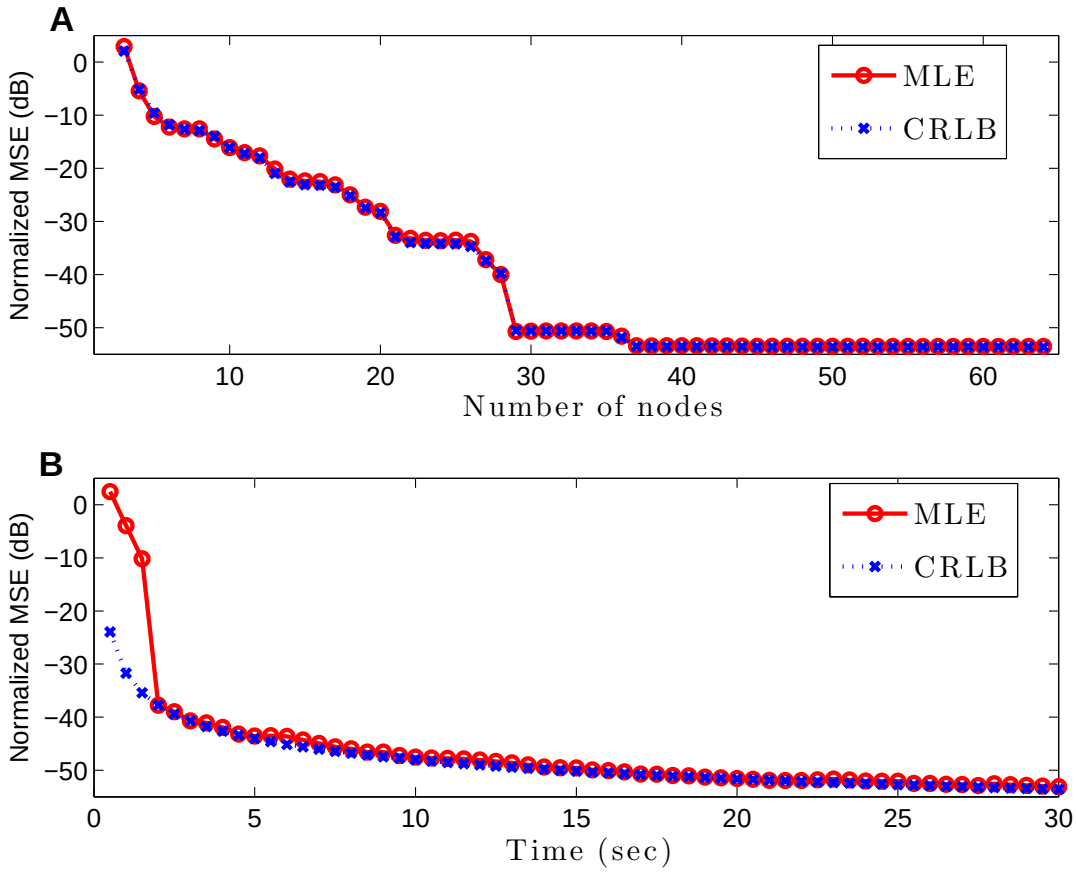



Figure 3.1: Normalized MSE and CRLB of the MLE as function of (a) number of nodes, and (b) time samples.

expect, the overall performance obtained from the BLUE estimator is not as good as that from the MLE due to the linear approximation applied on the observation model in (3.10). However, performance of the BLUE estimator based localization improves as the number of nodes and/or time samples increases. This is because for $N, T \to \infty$, the complementary error function in (3.3) tends to be equal to 1, causing the linearization having almost no effect on the approximation.

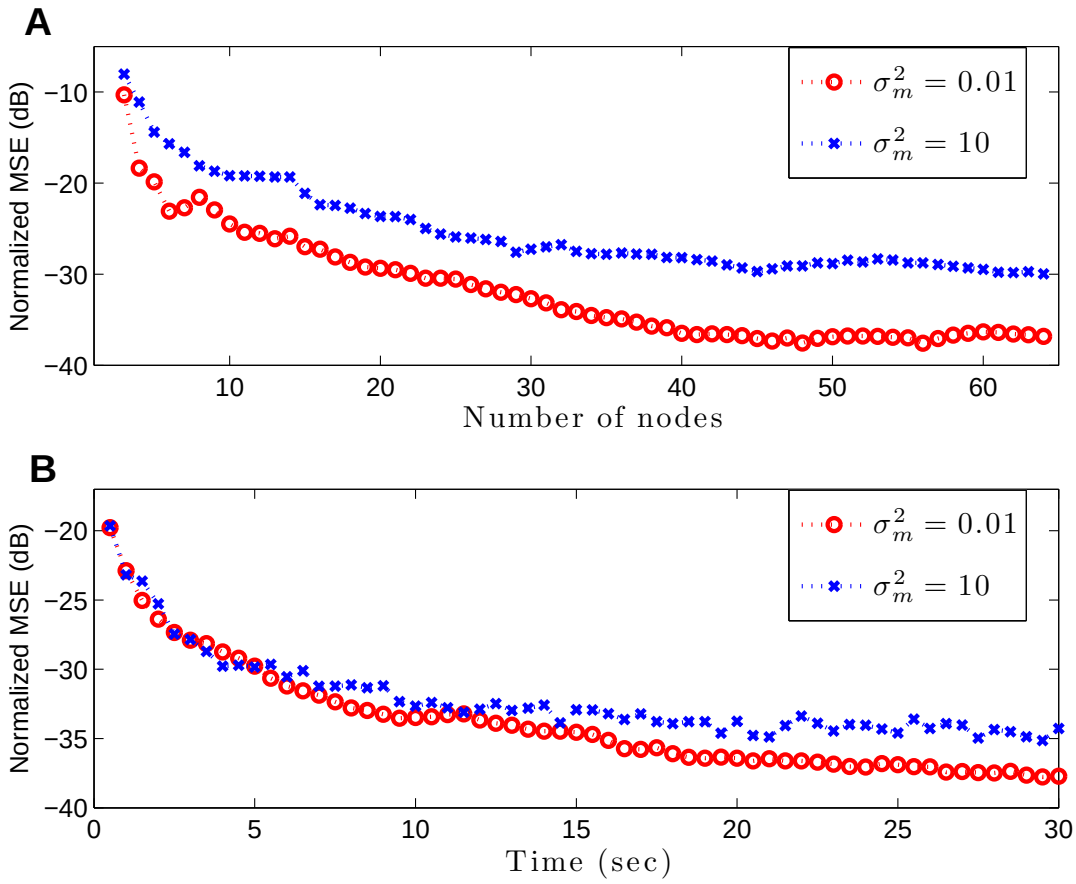


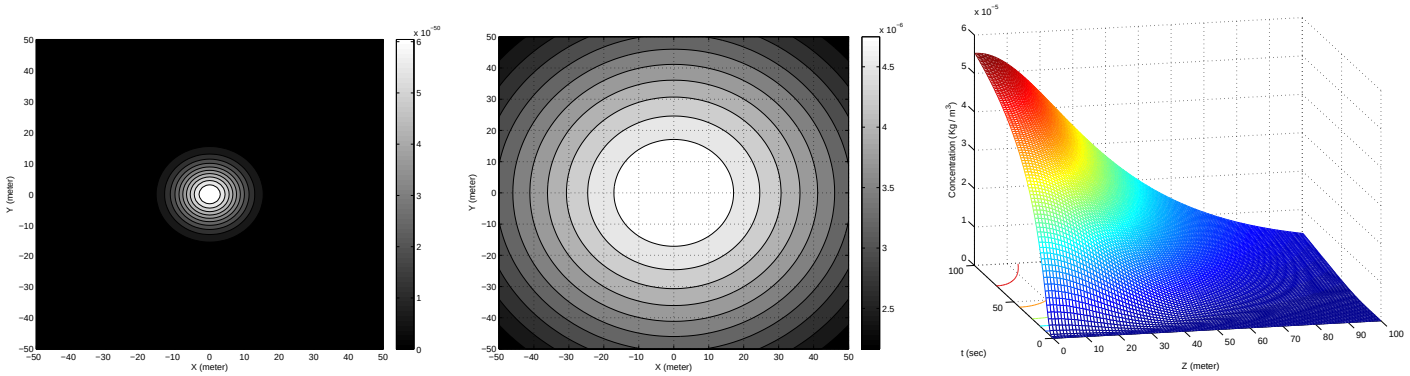

Figure 3.2: Normalized MSE of the BLUE as function of (a) number of nodes, and (b) time samples.

### 3.6.2 Reconstructed Spatio-Temporal Concentration Distribution

In this section, we reconstruct the space-time variance concentration distribution based on our estimated source location $\hat{\mathbf{r}}_0 = [\hat{x}_0, \hat{y}_0, \hat{z}_0]^T$ using MLE and BLUE, and compare them with the original concentration distribution. For better clarity of exposition, we consider a diffusive field of $100 \times 100 \times 100$ $m^3$. The rest of the parameters remain same as before.

Figure 3.3 shows the actual concentration distribution in space and in time. As it can be seen at time $t = 1$ sec, the diffusive source located at the center $\mathbf{r}_0 = [0, 0, 0]^T$ just started

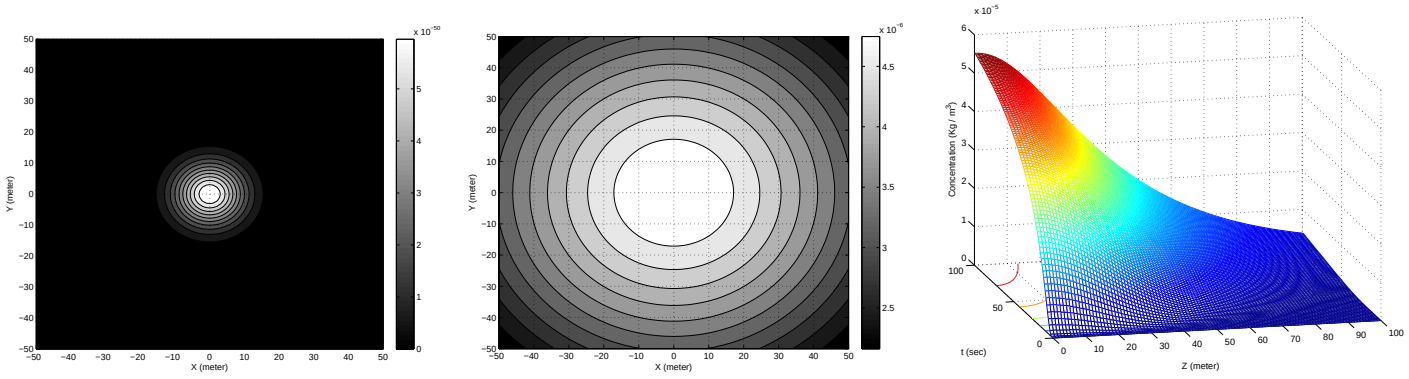
to diffuse chemical substance into the field. At time $t = 100$ sec, the chemical substance emitting from the source is diffused through the medium and spread out to larger distance from the source position. Figure 3.3(c) shows the change in concentration along the z-axis over time for x=25m and y=25m. It is interesting to see that the concentration in this case increases with the increase in time and decreases with the increase in distance from the source position.




(a) (b) (c)

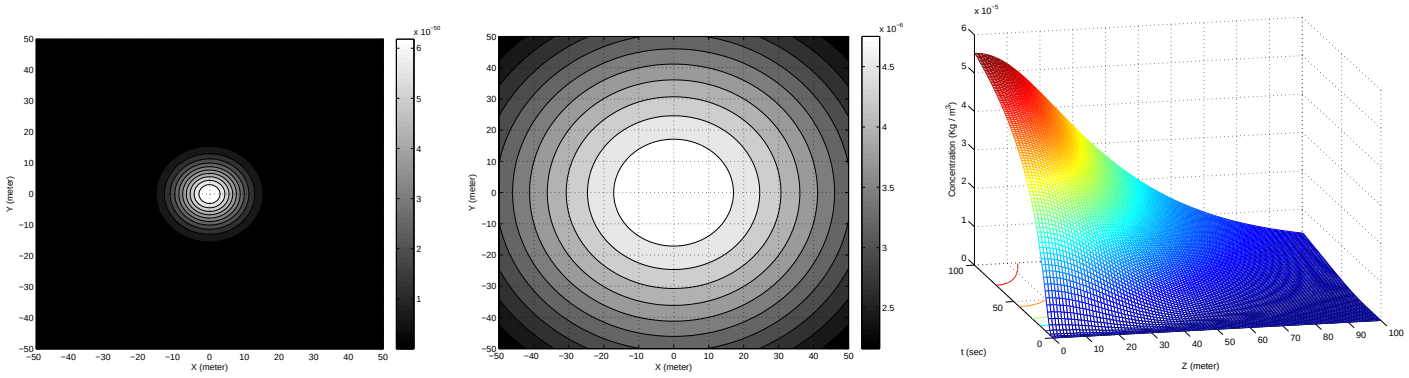
Figure 3.3: Concentration distribution in space (x-y-z coordinates) at times (a) $t = 1$ sec and (b) $t = 100$ sec for $z = 100$m, (c) Concentration distribution along z-axis and in time for x=25m and y=25m.

Figure 3.4 and 3.5 show the reconstructed concentration distribution for MLE and BLUE respectively. As it can be seen from both of these figures, though the concentration distribution pattern are same as that for the original, but there is slight shifting in the concentration distribution peak both over space and time because of the estimated source locations being different for both MLE and BLUE.




(a) (b) (c)

Figure 3.4: Concentration distribution in space (x-y-z coordinates) at times (a) $t = 1$ sec and (b) $t = 100$ sec for $z = 100$m, (c) Concentration distribution along z-axis and in time for x=25m and y=25m for MLE.




(a) (b) (c)

Figure 3.5: Concentration distribution in space (x-y-z coordinates) at times (a) $t = 1$ sec and (b) $t = 100$ sec for $z = 100$m, (c) Concentration distribution along z-axis and in time for x=25m and y=25m for BLUE.

## 3.7 Conclusion

In this chapter, we proposed two parametric estimation methods based on the MLE and BLUE for estimating static diffusive source location using sensor measurements. We also

obtained the Cramér-Rao lower bound as theoretical performance bound for source localization. While the MLE requires complete knowledge of the pdf, knowing only 1st and 2nd order information of data are enough for BLUE-based technique. On the other hand, BLUE can only be used for linear observation models. Since in our case, concentration is highly nonlinear in space and time, hence linearization is necessary before using BLUE for parameter estimation. However, the performance of the BLUE deteriorates due to linearization effect for lower number of sensing nodes and time samples.

It was observed that performance of the MLE is slightly better than the CRLB in some cases. This is because of the fact that the MLE in this case is biased, and therefore it can outperform the CRLB by trading variance for bias. As one would expect, the MLE performs better than the BLUE-based diffusive source localization method. However, the later shows satisfactory performance trend for large number of sensing nodes and time samples. This is because, linearization tends to have negligible effect on the performance of the BLUE as the number of samples increases. Finally, the estimated source locations were used to reconstruct the spatio-temporal concentration distributions, and the outputs were found to be satisfactory.

# Appendix 3A

## Proof of Theorem 1

To prove the consistency for the MLE of source localization, we follow the technique described in [40]. We show the proof w.r.t. $\theta_1 = x_0$ and it can be easily followed to prove the consistency for $\theta_2 = y_0$ and $\theta_3 = z_0$ without any loss of generality. we have to prove that

$$\lim_{N,T\to\infty} \frac{1}{d_{N,T}} \sum_{j=1}^{N} \sum_{k=1}^{T} K_{N,T}(x_0; x_0') \text{ exists, and} \tag{3.14}$$

$$\lim_{N,T\to\infty} \frac{1}{d_{N,T}^2} \sum_{j=1}^{N} \sum_{k=1}^{T} \left[ \frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial x_0} \right]^2 \;=\; 0, \tag{3.15}$$

for some sequence $\{d_{N,T} > 0\}_{N,T=1}^{\infty}$, where $\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial x_0}$ is defined by (3.8), and

$$K_{N,T}(x_0; x_0') \triangleq \sum_{j=1}^{N} \sum_{k=1}^{T} \left\{ \frac{\partial c_{jk}(\boldsymbol{\theta}')}{\partial x_0'} \right\} \left[ c_{jk}(\boldsymbol{\theta}) - c_{jk}\left(\boldsymbol{\theta}'\right) \right], \tag{3.16}$$

with $x_0 \neq x_0' \Rightarrow \mathbf{r}_0 \neq \mathbf{r}_0'$ and $\boldsymbol{\theta} \neq \boldsymbol{\theta}'$.

For the sake of brevity, while doing the proof for $x_0$, we denote $c_{jk}(\boldsymbol{\theta}) \triangleq c_{jk}(x_0)$ and $\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial x_0} \triangleq c_{jk}'(x_0)$. Since both $c_{jk}(x_0)$ and $c_{jk}'(x_0)$ are continuous functions of $x_0$. Hence, we have

$$\begin{aligned}
K_{NT}(x_0; x_0') \;&\leq\; \left| \sum_{j=1}^{N} \sum_{k=1}^{T} c_{jk}'\left(x_0'\right) \left[ c_{jk}(x_0) - c_{jk}\left(x_0'\right) \right] \right|, \\
&\leq\; \sum_{j=1}^{N} \sum_{k=1}^{T} \left| c_{jk}'\left(x_0'\right) \left[ c_{jk}(x_0) - c_{jk}\left(x_0'\right) \right] \right|. \\
&\leq\; \left[ \sum_{j=1}^{N} \sum_{k=1}^{T} \frac{\mu |x_j - x_0'|}{4\pi\kappa |\mathbf{r}_j - \mathbf{r}_0'|^2} + \frac{1}{\sqrt{\pi\kappa(t_j - t_I)}} \right] \\
&\quad \times \left[ \sum_{j=1}^{N} \sum_{k=1}^{T} \frac{\mu}{4\pi\kappa} \left\{ \frac{1}{|\mathbf{r}_j - \mathbf{r}_0|} + \frac{1}{|\mathbf{r}_j - \mathbf{r}_0'|} \right\} \right].
\end{aligned} \tag{3.17}$$

Since $t_I < t_k, \forall k$, we can have the following:

$$\begin{aligned}
\sum_{k=1}^{T} \frac{1}{\sqrt{t_k - t_I}} &< \sum_{k=1}^{T} \left| t_1^{-1/2} \right| \left| \left(1 - \frac{t_I}{t_k}\right)^{-1/2} \right| \\
&\leq \frac{1}{\sqrt{t_1}} \left[ T \max_{1 \leq k \leq T} \left\{ \left(1 - \frac{t_I}{t_k}\right)^{-1/2} \right\} \right] \\
&\leq \underbrace{T \left[ \frac{1}{\sqrt{t_1}} \max_{1 \leq k \leq T} \left\{ \left(1 - \frac{t_I}{t_k}\right)^{-1/2} \right\} \right]}_{S} \\
\therefore \sum_{k=1}^{T} \frac{1}{\sqrt{t_k - t_I}} &< TS. \qquad (3.18)
\end{aligned}$$

Similarly, for $\left|\frac{t_I}{t_k}\right| < 1$, we have $\sum_{k=1}^{T} \frac{1}{t_k - t_I} < TQ_1$, where $S$ and $Q_1$ are positive real values.

For practical consideration, assuming $0 \leq \frac{|x_j - x_0'|}{|\mathbf{r}_j - \mathbf{r}_0|^2} \leq P$, $0 < \frac{1}{|\mathbf{r}_j - \mathbf{r}_0|} \leq M_1$, $0 < \frac{1}{|\mathbf{r}_j - \mathbf{r}_0'|} \leq M_2$ and using (3.18), $K_{NT}\left(x_0; x_0'\right)$ can be written as:

$$K_{NT}\left(x_0; x_0'\right) \leq \left[ \frac{\mu P}{4\pi\kappa} + \frac{S}{\sqrt{\pi\kappa}} \right] \left[ \frac{\mu}{4\pi\kappa} (M_1 + M_2) \right] \times N^2 T^2 \qquad (3.19)$$

If $d_{N,T} = N^3 T^3 > 0$ for $N \geq 1, T \geq 1$, then we can claim that
$\lim_{N,T \to \infty} \frac{1}{d_{N,T}} \sum_{j=1}^{N} \sum_{k=1}^{T} K_{N,T}(x_0; x_0')$ exists.

Now, we need to prove that $\lim_{N,T \to \infty} \frac{1}{d_{N,T}^2} \sum_{j=1}^{N} \sum_{k=1}^{T} \left\{ c_{jk}'(x_0) \right\}^2 = 0$. Therefore, from

(3.8), we can write:

$$\begin{aligned}\left\{c'_{jk}(x_0)\right\}^2 &= \frac{\mu^2|x_j-x_0|^2}{16\pi^2\kappa^2|\mathbf{r}_j-\mathbf{r}_0|^2}\left[\frac{\operatorname{erfc}^2\left\{\frac{|\mathbf{r}_j-\mathbf{r}_0|}{2\sqrt{\kappa(t_k-t_I)}}\right\}}{|\mathbf{r}_j-\mathbf{r}_0|^2}+\frac{\exp\left\{\frac{-|\mathbf{r}_j-\mathbf{r}_0|^2}{2\kappa(t_k-t_I)}\right\}}{\pi\kappa(t_k-t_I)}\right.\\ &\quad\left.+\frac{2\operatorname{erfc}\left\{\frac{|\mathbf{r}_j-\mathbf{r}_0|}{2\sqrt{\kappa(t_k-t_I)}}\right\}\exp\left\{\frac{-|\mathbf{r}_j-\mathbf{r}_0|^2}{4\kappa(t_k-t_I)}\right\}}{|\mathbf{r}_j-\mathbf{r}_0|\sqrt{\pi\kappa(t_k-t_I)}}\right]\\ &\leq \frac{\mu^2}{16\pi^2\kappa^2}\frac{|x_j-x_0|^2}{|\mathbf{r}_j-\mathbf{r}_0|^2}\left[\frac{1}{|\mathbf{r}_j-\mathbf{r}_0|^2}+\frac{1}{\pi\kappa(t_k-t_I)}+\frac{2}{|\mathbf{r}_j-\mathbf{r}_0|\sqrt{\pi\kappa(t_k-t_I)}}\right].\end{aligned} \tag{3.20}$$

Assuming $0 \leq \frac{|x_j-x_0|^2}{|\mathbf{r}_j-\mathbf{r}_0|^2} \leq Q_2$ and using $\sum_{k=1}^{T}\frac{1}{t_k-t_I} < TQ_1$, we obtain the following from (3.20),

$$\begin{aligned}\left\{c'_{jk}(x_0)\right\}^2 &< \frac{\mu^2Q_2}{16\pi^2\kappa^2}\left[M_1^2+\frac{1}{\pi\kappa(t_k-t_I)}+\frac{2M_1}{\sqrt{\pi\kappa(t_k-t_I)}}\right]\\ \therefore\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{c'_{jk}(x_0)\right\}^2 &< \frac{\mu^2Q_2}{16\pi^2\kappa^2}\left[\sum_{j=1}^{N}\sum_{k=1}^{T}M_1^2+\frac{1}{\pi\kappa}\sum_{j=1}^{N}\sum_{k=1}^{T}\frac{1}{t_k-t_I}+\frac{2M_1}{\sqrt{\pi\kappa}}\sum_{j=1}^{N}\sum_{k=1}^{T}\frac{1}{\sqrt{t_k-t_I}}\right]\\ &< \frac{\mu^2Q_2}{16\pi^2\kappa^2}\left[M_1^2NT+\frac{Q_1}{\pi\kappa}NT+\frac{2M_1S}{\sqrt{\pi\kappa}}NT\right]\\ &= \frac{\mu^2Q_2}{16\pi^2\kappa^2}\left[M_1^2+\frac{Q_1}{\pi\kappa}+\frac{2M_1S}{\sqrt{\pi\kappa}}\right]\times NT.\end{aligned} \tag{3.21}$$

For $d_{N,T}=N^3T^3>0, N\geq 1, T\geq 1$, we have

$$\lim_{N,T\to\infty}\frac{1}{d_{N,T}^2}\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{c'_{jk}(x_0)\right\}^2 = 0. \tag{3.22}$$

Similarly, for $y_0$ and $z_0$, we can also claim that the MLE to the diffusive source localization problem is consistent when the number of sensor nodes and time samples go to infinity. ■

# Appendix 3B

## Proof of Theorem 2

To prove the asymptotic normality of the the MLE, let us first define the following notations, which will be useful in presenting the sufficient conditions for asymptotic normality. For $p\left(y_{jk}|\boldsymbol{\theta}\right)$, we define:

$$\Phi_{jk}\left(y_{jk}|\boldsymbol{\theta}\right) \;=\; \log p\left(y_{jk}|\boldsymbol{\theta}\right).$$

Let $\dot{\boldsymbol{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)$ be the $3\times 1$ vector whose $u$-th component is

$$\dot{\boldsymbol{\Phi}}_{jk,u}\left(y_{jk}|\boldsymbol{\theta}\right) \;=\; \frac{\partial}{\partial\theta_u}\left\{\Phi_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)\right\},$$

and let $\ddot{\boldsymbol{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)$ be the $3\times 3$ matrix whose $(u,v)$-th component is

$$\ddot{\boldsymbol{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right) \;=\; \frac{\partial^2}{\partial\theta_u\partial\theta_v}\left\{\Phi_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)\right\}.$$

It is proved in [69] that the ML estimates for the independent not identically distributed case are asymptotically normal if the following nine conditions are satisfied:

C1. The parameter space is an open subset of $\mathbf{R}^3$.

C2. The ML estimator is consistent.

C3. $\dot{\boldsymbol{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)$ and $\ddot{\boldsymbol{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right)$ exist almost surely.

C4. $\ddot{\boldsymbol{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)$ is a uniformly continuous function of $\boldsymbol{\theta}$ in $j$ and $k$ almost surely and is a measurable function of $y_{jk}$.

C5. $\mathbb{E}\left[\dot{\boldsymbol{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)\right]=0,\forall j,k.$

C6. $\mathbb{E}\left[\dot{\mathbf{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)\dot{\mathbf{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)^{T}\right] = -\mathbb{E}\left[\ddot{\mathbf{\Phi}}_{jk}\left(y_{jk}|\boldsymbol{\theta}\right)\right]$.

C7. $\bar{I}_{\boldsymbol{\theta}} = \lim_{N,T\to\infty}\frac{1}{NT}I_{\boldsymbol{\theta}}$ is positive definite for all $N, T \geq 1$.

C8. $\mathbb{E}\left[\left|\dot{\mathbf{\Phi}}_{jk,u}\left(y_{jk}|\boldsymbol{\theta}\right)\right|^{3}\right] \leq K$ for all $j, k$ and $1 \leq u \leq 3$, where $K$ is some finite real number.

C9. There exist $\epsilon > 0$ and random variables $B_{jk,u,v}(y_{jk})$ such that

i) $\sup\left\{\left|\ddot{\mathbf{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right)\right| : ||\boldsymbol{\theta} - \boldsymbol{\theta}_0|| \leq \epsilon\right\} \leq B_{jk,u,v}(y_{jk})$, where $\boldsymbol{\theta}_0$ is the actual parameter vector.

ii) $\mathbb{E}\left[|B_{jk,u,v}(y_{jk})|^{2}\right] \leq K$ for some real finite number $K$.

Below, we verify the necessary conditions mentioned in [69] for our obtained MLE to be asymptotically normal.

C1. From practical point of view, there is no loss in generality in assuming that $\boldsymbol{\theta}_0 \in \mathring{\Lambda}$, where $\mathring{\Lambda} \subset \Lambda$ is an open subset of $\Lambda$. Thus C1 is satisfied.

C2. Since from Appendix 3A, MLE of diffusive source location is consistent, hence it is also consistent even when $\boldsymbol{\theta}_0 \in \mathring{\Lambda} \subset \Lambda$. Therefore C2 is satisfied.

C3. From the notations defined above, we have

$$\dot{\mathbf{\Phi}}_{jk,u}\left(y_{jk}|\boldsymbol{\theta}\right) = \frac{1}{\sigma^2}\left(y_{jk} - c_{jk}(\boldsymbol{\theta}) - b\right)\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial \theta_u}, \tag{3.23}$$

where $\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial \theta_u}$ is given by (3.8) for $u = 1, 2, 3$ respectively, and

$$\ddot{\mathbf{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right) = \frac{1}{\sigma^2}\left[\left(y_{jk} - c_{jk}(\boldsymbol{\theta}) - b\right)\frac{\partial^2 c_{jk}(\boldsymbol{\theta})}{\partial \theta_u \partial \theta_v} - \frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial \theta_u}\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial \theta_v}\right]. \tag{3.24}$$

It can be easily verified that $\frac{\partial^2 c_{jk}(\boldsymbol{\theta})}{\partial \theta_u \partial \theta_v}$ exists (omitted for tedious and lengthy calculation). Therefore both $\dot{\mathbf{\Phi}}_{jk,u}\left(y_{jk}|\boldsymbol{\theta}\right)$ and $\ddot{\mathbf{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right)$ exist almost surely.

C4. It can be easily verified from (3.24) that $\ddot{\mathbf{\Phi}}_{jk}(y_{jk}|\boldsymbol{\theta})$ is a continuous function of $\boldsymbol{\theta}$ in $j$ and $k$. Because of $\boldsymbol{\theta} \in \Lambda$, and $\ddot{\mathbf{\Phi}}_{jk}(y_{jk}|\boldsymbol{\theta})$ being a continuous mapping of $\boldsymbol{\theta}$, therefore using Theorem 4.19 in [67], we can claim that $\ddot{\mathbf{\Phi}}_{jk}(y_{jk}|\boldsymbol{\theta})$ is indeed uniformly continuous on $\boldsymbol{\theta}$ in $j$ and $k$.

Since $\ddot{\mathbf{\Phi}}_{jk}(y_{jk}|\boldsymbol{\theta}) : y_{jk} \rightarrow \mathbb{R}$ is a continuous function of $y_{jk}$, with $y_{jk} \in \mathbb{R}$ being a measurable space, hence $\ddot{\mathbf{\Phi}}_{jk}(y_{jk}|\boldsymbol{\theta})$ is also a measurable function of $y_{jk}$ [67].

C5. From (3.23), we obtain

$$\mathbb{E}\left[\dot{\mathbf{\Phi}}_{jk,u}(y_{jk}|\boldsymbol{\theta})\right] = \frac{1}{\sigma^2}\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial \theta_u}\mathbb{E}\left[(y_{jk} - c_{jk}(\boldsymbol{\theta}) - b)\right] = 0, \forall j, k \text{ and } u.$$

C6. As given in [69], C6 is implied if the following is true:

$$\frac{\partial^2}{\partial\theta_u \partial\theta_v}\int p(y_{jk}|\boldsymbol{\theta})dy_{jk} = \int \frac{\partial^2}{\partial\theta_u\partial\theta_v}p(y_{jk}|\boldsymbol{\theta})dy_{jk}, \forall j, k, u \text{ and } v. \tag{3.25}$$

Since in our case $p(y_{jk}|\boldsymbol{\theta}) \sim \mathcal{N}(c_{jk}(\boldsymbol{\theta}) + b, \sigma^2)$, and $p(y_{jk}|\boldsymbol{\theta})$ is continuous function of $y_{j,k}$, hence (3.25) is valid and thus C6 is satisfied.

C7. The Fisher information matrix (FIM) in our case is a $3\times3$ matrix with $(u, v)$-th element given by:

$$[I_{\boldsymbol{\theta}}]_{u,v} = \frac{1}{\sigma^2}\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\}.$$

Using the Cauchy-Schwarz inequality,

$$\left(\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}^2\right)\left(\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\}^2\right) - \left(\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\}\right)^2 \geq 0,$$

with equality if and only if $\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}$ and $\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\}$ are linearly dependent, which is impossible as the coordinates of the locations of the distributed sensor nodes are

linearly independent. Therefore, all the leading principle minors of $I_{\boldsymbol{\theta}}$ can be shown to be positive and hence the FIM for fixed $N$ and $T$ is a positive-definite matrix.

From the prove of consistency of MLE in Appendix 3A, it can be observed that $\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}^2$ and $\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\}$ exist and are bounded for all $u, v$. Hence the $(u, v)$-th element of $\bar{I}_{\boldsymbol{\theta}}$,

$$\left[\bar{I}_{\boldsymbol{\theta}}\right]_{u,v} = \frac{1}{\sigma^2}\lim_{N,T\to\infty}\frac{1}{NT}\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right\}\left\{\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right\},$$

exits for all $u, v$. Therefore, using the same reasoning used for proving the positive-definiteness of the FIM for the finite case, we can claim that $\bar{I}_{\boldsymbol{\theta}}$ is also positive-definite.

C8. From (3.23), we have the following:

$$\mathbb{E}\left[\left|\dot{\boldsymbol{\Phi}}_{jk,u}(y_{jk}|\boldsymbol{\theta})\right|^3\right] \leq \frac{1}{\sigma^6}\left|\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right|^3\mathbb{E}\left[|y_{jk}-c_{jk}(\boldsymbol{\theta})-b|^3\right].$$

As we know, for $X\sim\mathcal{N}(\mu,\sigma^2)$,

$$\mathbb{E}\left[|X-\mu|^p\right] = \sigma^p\frac{2^{\frac{p}{2}}\Gamma(\frac{p+1}{2})}{\sqrt{\pi}},$$

therefore

$$\mathbb{E}\left[\left|\dot{\boldsymbol{\Phi}}_{jk,u}(y_{jk}|\boldsymbol{\theta})\right|^3\right] \leq \frac{2}{\sigma^2}\sqrt{\frac{2}{\pi}}\left|\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right|^3 \leq K.$$

Hence, C8 is satisfied $\forall j, k,$ and $u$.

C9. Since $\ddot{\boldsymbol{\Phi}}_{jk,u,v}(y_{jk}|\boldsymbol{\theta})$ is a uniformly continuous function of $\boldsymbol{\theta}$ (shown in C4), hence for each $\delta>0$, there exists one $\epsilon>0$ such that,

$$\left|\ddot{\boldsymbol{\Phi}}_{jk,u,v}(y_{jk}|\boldsymbol{\theta})-\ddot{\boldsymbol{\Phi}}_{jk,u,v}(y_{jk}|\boldsymbol{\theta}_0)\right| < \delta \qquad \forall\,||\boldsymbol{\theta}-\boldsymbol{\theta}_0||<\epsilon.$$

From (3.24), we have

$$\left|\ddot{\boldsymbol{\Phi}}_{jk,u,v}(y_{jk}|\boldsymbol{\theta})\right| \leq \frac{1}{\sigma^2}\left[|e_{jk}|\left|\frac{\partial^2 c_{jk}(\boldsymbol{\theta})}{\partial\theta_u\partial\theta_v}\right|+\left|\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_u}\right|\left|\frac{\partial c_{jk}(\boldsymbol{\theta})}{\partial\theta_v}\right|\right].$$

Therefore, for each $\epsilon > 0$, the following can be obtained

$$\begin{aligned}\sup\left\{\left|\ddot{\mathbf{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right)\right|\right\} &< \delta + \sup\left\{\left|\ddot{\mathbf{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}_0\right)\right|\right\} \\ &\leq \delta + \frac{1}{\sigma^2}\left[|e_{jk}|\sup\left\{\left|\frac{\partial^2 c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_u\partial\theta_v}\right|\right\}\right. \\ &\quad \left.+\sup\left\{\left|\frac{\partial c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_u}\right|\right\}\sup\left\{\left|\frac{\partial c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_v}\right|\right\}\right].\end{aligned}$$

Since $\frac{\partial c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_u}$, $\frac{\partial c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_v}$ and $\frac{\partial^2 c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_u\partial\theta_v}$ are continuous functions of $\boldsymbol{\theta}_0 \in \mathring{\Lambda} \subset \Lambda, \forall j,k,u$ and $v$, hence $\sup\left\{\left|\frac{\partial c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_u}\right|\right\}$, $\sup\left\{\left|\frac{\partial c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_v}\right|\right\}$ and $\sup\left\{\left|\frac{\partial^2 c_{jk}(\boldsymbol{\theta}_0)}{\partial\theta_u\partial\theta_v}\right|\right\}$ can be bounded replaced by finite real numbers. Therefore,

$$\sup\left\{\left|\ddot{\mathbf{\Phi}}_{jk,u,v}\left(y_{jk}|\boldsymbol{\theta}\right)\right|\right\} \leq \delta + \left[|e_{jk}|\,K^{\sup}_{jk,u,v} + K^{\sup}_{jk,u}K^{\sup}_{jk,u}\right] = B_{jk,u,v}(e_{jk}),$$

where $K^{\sup}_{jk,u,v}$, $K^{\sup}_{jk,u}$ and $K^{\sup}_{jk,u}$ are some finite real numbers and $B_{jk,u,v}(e_{jk})$ are random variables. Since $\mathbb{E}\left\{|e_{jk}|\right\} = \sigma\sqrt{\frac{2}{\pi}}$ and $\mathbb{E}\left\{|e_{jk}|^2\right\} = \sigma^2$, hence

$$\mathbb{E}\left[|B_{jk,u,v}(e_{jk})|^2\right] \leq K,$$

where $K$ is a finite real number.

Therefore, the obtained MLE of the diffusive source location is asymptotically normal when the number of sensor nodes and time samples go to infinity.■

# Chapter 4

# Dirichlet Process Mixture Model Based Diffusive Source Localization

## 4.1 Introduction

In this chapter, we consider the problem of diffusive source localization using nonparametric estimation approach. In most cases, accuracy of the widely used parametric estimation techniques depends on the sophistication of underlying statistical model, and priors. Since real data are often complicated, hence accurate and suitable parametric representation is sometimes difficult to obtain. As an alternative, we explore the rich tools of machine learning, which offers efficient learning algorithms or techniques allowing the WSN to detect threats intelligently. Since the diffusive field can be very dynamic and intelligent hazardous source could pose severe threat, sticking to a fixed statistical model might not be a smart idea. To that end, machine learning can be a suitable approach to deal with those dynamic or intelligent threats. To the best of our knowledge, though there are many literature on the application of machine learning in WSN area [70–73], no efforts have been made to utilize the rich collection of machine learning tools in diffusive source detection and estimation. We

plan to consider "*Smart WSN*" with cognition and machine learning abilities that is capable of adapting its actions in order to reach a desired quality of detection and/or estimation about diffusive source.

In this chapter, to overcome the aforementioned drawbacks of parametric estimation and localization, we propose to use nonparametric Bayesian methods, in particular Dirichlet process mixture model (DPMM) which provides us with a flexible and data-driven estimation process to estimate the location of diffusive source from sensor measurements [59]. DPMM has been previously applied in various pragmatic problems, such as, classification [74], image segmentation [75] and language modeling [76]. The nonparametric nature of DPMM framework allows for minimal knowledge of measurement data, and can estimate unknown parameters of corresponding probability distribution from data itself in order to best fit the underlying the observation model. We propose DPMM classifier based algorithm to estimate the location of continuous diffusive source emitting substance in a diffusive environment, and provide analytical proof of convergence. The proposed algorithm is also extended to the scenario when multiple diffusive sources of same kind are present in the diffusive field of interest.

The remainder of this chapter is organized as follows: Section 4.2 discusses physical modeling of a diffusion phenomena and measurement model for static diffusive source localization using sensor network. Brief descriptions on DPMM and Gibbs Sampling [77] are given in section 4.3. The proposed DPMM-based algorithm for single static diffusive source localization, along with analytical proof of convergence is discussed in section 4.4. The proposed algorithm is extended to multiple diffusive source localizations in section 4.5. Simulation results for the proposed algorithms are shown in section 4.6. Finally, section 4.7 concludes the chapter by summarizing our results.

## 4.2 Physical and Measurement Models

In this section, we use the same physical and measurement models as discussed in section 3.2 and 3.3 of Chapter 3, where the substance concentration and observation for a sensing node at location $\mathbf{r}_j$ and at time $t_k$ are defined as follows:

$$c(\mathbf{r}_j, t_k) = \frac{\mu}{4\pi\kappa|\mathbf{r}_j - \mathbf{r}_0|}\text{erfc}\left(\frac{|\mathbf{r}_j - \mathbf{r}_0|}{2\sqrt{\kappa(t_k - t_I)}}\right), \tag{4.1}$$

and

$$y(\mathbf{r}_j, t_k) = c(\mathbf{r}_j, t_k) + e(\mathbf{r}_j, t_k) + b, \tag{4.2}$$

where $c(\mathbf{r}_j, t_k)$ is the concentration of interest, $t_I$ is initial substance release time, $\kappa$ is the diffusivity constant, $\mathbf{r}_0$ is the actual source location, $b$ is a bias term, and $e(\mathbf{r}_j, t_k) \sim \mathcal{N}(0, \sigma^2)$ is the sensor noise assumed to be independent in both time and space. For the sake of brevity, (4.2) can be rewritten in the simplified form as:

$$y_{j,k} = c_{j,k}(\boldsymbol{\theta}) + e_{j,k} + b, \tag{4.3}$$

where $y_{j,k} = y(\mathbf{r}_j, t_k)$, $e_{j,k} = e(\mathbf{r}_j, t_k)$ and $c_{j,k}(\boldsymbol{\theta}) = c(\mathbf{r}_j, t_k)$, $\boldsymbol{\theta} \in \mathbb{R}^{n\times 1}$ is the unknown source and medium parameter vector that we are interested to estimate. For our static diffusive source localization problem, only $[x_0, y_0, z_0]^T$ are the parameters of interest.

## 4.3 Learning-based Diffusive Source Estimation

One of the main ideas of learning is to *explore* the environment to get evaluative feedback and *exploit* the resources to increase the rewards. Each time an agent executes action under certain environment conditions, it gets either a *good* or a *bad* feedback. In our case, at each

time instant, the WSN has an estimate about the source parameters based on the sensor observations. This estimate is in general different from the actual parameters due to noise. The goal of the WSN is to reach close to the actual solution by adapting its actions. Below we provide a brief description of DPMM, which will be used as the backbone of our proposed nonparametric diffusive source localization algorithms.

### 4.3.1 Dirichlet Process

A Dirichlet process $DP(\alpha_0, G_0)$ is defined to be the distribution of a random probability measure $G$ over a measurable space $(\Theta, \mathcal{B})$, such that, for any finite measurable partition $(A_1, A_2, \ldots, A_r)$ of $\Theta$, the random vector $(G(A_1), G(A_2), \ldots, G(A_r))$ is distributed as a finite dimensional Dirichlet distribution with parameters $(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_r))$, for $\alpha_0 > 0$. It can be denoted as follows [59]:

$$(G(A_1), G(A_2), \ldots, G(A_r)) \sim Dir\left(\alpha_0 G_0(A_1), \alpha_0 G_0(A_2), \ldots, \alpha_0 G_0(A_r)\right).$$

$G \sim DP(\alpha_0, G_0)$ represents the probability measure $G$ that is drawn from the Dirichlet process $DP(\alpha_0, G_0)$. That is, $G$ is a random probability measure whose distribution is given by the Dirichlet process $DP(\alpha_0, G_0)$.

The first and direct approach of constructing the random probability distribution $G$ is the *stick-breaking* method [59]. Another way to construct a Dirichlet process is to characterize the distribution of the drawings $\theta$ of $G$, given a certain realization $G$ of $DP(\alpha_0, G_0)$. It is to be noted that $G$ is discrete with probability 1. The infinite discrete support for $G$ makes it a suitable candidate for non-parametric Bayesian methods. However, since this method involves infinite sum in $G$ for practical implementation, other alternatives methods to construct $G$ are Polya urn model and the Chinese Restaurant Process (CRP) [59, 78]. In case of a Hierarchical Dirichlet process (HDP) [59], $G_0$ itself is drawn from a Dirichlet process, such as $G_0 \sim DP(\gamma, H)$.

### 4.3.2 Dirichlet Process Mixture Model (DPMM)

Dirichlet process mixture model is a suitable candidate for non-parametric classification problems. DPMM has a non-parametric prior on the parameters of the mixture model [59]. A DPMM is defined as follows:

$$\left.\begin{array}{rcl} G & \sim & DP(\alpha_0, G_0) \\ \theta_i|G & \sim & G \\ y_i|\theta_i & \sim & f(\theta_i) \end{array}\right\}, \tag{4.4}$$

where, $\theta_1, \theta_2, \ldots$ are i.i.d. random variables distributed according to $G$ and $y_i$'s are the observations. In this chapter, we will exploit the excellent classification capability of DPMM to perform diffusive source localization.

### 4.3.3 Gibbs Sampling

Let us consider a sequence of observations $\boldsymbol{y}_{1:N} \triangleq \{y_j\}_{j=1}^{N}$ and assume that these observations are drawn from an unknown mixture model. If we do not have enough information except for the sensor measurements and the family of distributions they belong to, it is reasonable to assume a non-parametric model, such as DPMM. To that end, let us assume that the mixture components $\theta_i$ are drawn from $G \sim DP(\alpha_0, G_0)$, for $G = \sum_{l=1}^{\infty} \pi_l \delta_{\phi_l}$, where $\delta_{\phi_l}$ are the unique values of $\theta_j$ and $\pi_l$ their corresponding probabilities. The problem is to estimate the mixture component $\hat{\theta}_j$ for each observation $y_j$. It can be done by applying the Gibbs sampling method proposed in [77]. In this particular method, the estimates $\hat{\theta}_j$ are sampled from the posterior distribution of $\theta_j$ given the observation sequence $\boldsymbol{y}_{1:N}$. Gibbs sampler (or Markov chain Monte Carlo (MCMC)) method has been described in detail in [79]. It is to be noted that $\theta_j$ can be generalized to multivariate case (in our case, we model source location vector $\boldsymbol{\theta}$ accordingly).

In a DPMM, the posterior distribution of $\theta_j|\{\theta_i\}_{i\neq j}, \boldsymbol{y}_{1:N}$ can be computed as [77]:

$$\theta_j|\{\theta_i\}_{i\neq j}, \boldsymbol{y}_{1:N} \begin{cases} = \theta_i & \text{with Pr.} \quad q_i \\ \sim f(\theta_j|y_j) & \text{with Pr.} \quad q_0 \end{cases}, \tag{4.5}$$

where $q_i = \frac{f_{\theta_i}(y_j)}{\alpha_0 f(y_j)+\sum_{i=1,i\neq j}^{N} f_{\theta_i}(y_j)}$, $q_0 = \frac{\alpha_0 f(y_j)}{\alpha_0 f(y_j)+\sum_{i=1,i\neq j}^{N} f_{\theta_i}(y_j)}$, and $f(y_j) = \int_\theta f_\theta(y_j) G_0(\theta) d\theta$ is the marginal distribution of $y_j$, assuming a prior $G_0(\theta)$. Note that, the posterior distribution $f(\theta_j|y_j)$ can easily be obtained if $\theta_j$ has a *conjugate prior* for the likelihood $f_{\theta_j}(y_j)$. In this case, $G_0(\theta_j)$ and $f(\theta_j|y_j)$ will belong to the same family of distributions. In particular, if both the prior distribution $G_0(\theta_j)$ and the likelihood function $f_{\theta_j}(y_j)$ are Gaussian, then the posterior distribution $f(\theta_j|y_j)$ will also be Gaussian. Therefore, almost all of the DPMM problems assume conjugate priors in their formulations [76, 77]. If the sequence of observations $\boldsymbol{y}_{1:N}$ drawn from a DPMM are normally distributed, given mixture component of parameters $\{\theta_j\}_{j=1}^{N}$ and if $y_j \in \mathbb{R}$, we let $y_j \sim \mathcal{N}(f(\mu_j), V_j)$, where $\theta_j = \mu_j$ for $j \in \{1, 2, \ldots, N\}$.

## 4.4 DPMM-based Single Diffusive Source Localization

Based on the discussions above, the following algorithm provides a recipe for estimating unknown source location of static diffusive source using DPMM. Let us denote by $\hat{\boldsymbol{\theta}}_{j,k}$, the estimated location of source by $j$-th sensor node at $k$-th time instant.

**Algorithm 1** `DPMM-based single diffusive source localization`

**STEP 1**: Take measurements $y_{j,k}, \forall j$ at $k$-th time instant.

**STEP 2**: Sample $\hat{c}_{j,k}$ from the posterior $p\left(c_{j,k}|y_{j,k}\right)$ and compute rough estimate $\hat{d}_{j,k}$, where $\hat{c}_{j,k}(\boldsymbol{\theta}) = \frac{\mu}{4\pi\kappa\hat{d}_{j,k}(\boldsymbol{\theta})}\text{erfc}\left[\frac{\hat{d}_{j,k}(\boldsymbol{\theta})}{2\sqrt{\kappa(t_k - t_I)}}\right]$ with $\hat{d}_{j,k}$ being the estimated distance between $j$-th node and the source.

**STEP 3**:

**while** Convergence conditions not satisfied **do**

**for** j= shuffle {1,2,...,N} **do**

Obtain $\hat{\boldsymbol{\theta}}_{j,k} \in \mathbb{R}^{n\times 1}$ using Gibbs sampling, where

$$\hat{\boldsymbol{\theta}}_{j,k} \left| \left\{\hat{\boldsymbol{\theta}}_{i,k}\right\}_{i\neq j}, \boldsymbol{y}_{1:N} \right. \sim q_0 \mathbf{G}_{j,k}(\hat{\boldsymbol{\theta}}_{j,k}) + \sum_{i=1, i\neq j}^{N} q_i \delta_{\hat{\boldsymbol{\theta}}_{i,k}}, \tag{4.6}$$

where $\mathbf{G}_{j,k}$ is a distribution such that $\hat{\boldsymbol{\theta}}_{j,k} \sim \mathbf{G}_{j,k}\left(\hat{\boldsymbol{\theta}}_{j,k}\right)$ and $\hat{\boldsymbol{\theta}}_{j,k} = \mathbf{r}_{j,k} + \hat{d}_{j,k}\angle\psi$, $\psi \sim \mathcal{U}[0, 2\pi]$, with weights $q_0$ and $q_i$ are defined as [77]:

$$q_0 \propto \frac{\alpha_0 c(s)}{\sqrt{L}}\left(1 + \frac{(y_j - m)^2}{sL}\right)^{-(1+s)/2} \tag{4.7}$$

$$q_i \propto \frac{1}{\sqrt{2V_i}} e^{-\frac{-(y_i - f(\boldsymbol{\mu}_i))^2}{2V_i}}, \tag{4.8}$$

for $i \in \{1, 2, \ldots, N\}$, $i \neq j$ and subject to $\sum_{i=1, i\neq j} q_i = 1$, with $L = \frac{(1+\tau)S}{s}$ and $c(s) = \Gamma\left(\frac{1+s}{2}\right)\Gamma\left(\frac{s}{2}\right)\sqrt{s}$.

**end for**

**end while**

**STEP 4**: Choose the cluster with maximum associated probability.

### 4.4.1 Description of the proposed Algorithm 1

Gibbs sampling described in section 4.3 provides us with the basic ingredient to Algorithm 1. Below we describe the proposed DPMM-based algorithm for single static diffusive source localization:

(1) In Step 1, the distributed sensor nodes takes concentration measurements from the diffusive field of interest, send the data to the fusion center (FC), and go back to sleep mode.

(2) In Step 2, rough estimate of of the source distance from a particular node is obtained using either numerical calculation or offline iso-concentration map (contour mapping) at the FC. The offline iso-concentration can be stored either in the distributed sensing nodes or at the FC to reduce computational cost during the estimation process.

It is to be noted that when the variance of the observations is not known, then we can set $\hat{c}_{j,k}(\boldsymbol{\theta}) \approx y_{j,k}, \forall j,k$ and obtain a crude estimate of $\hat{d}_{j,k}$ using contour mapping without any need for explicit pdf function of the measurement.

(3) In Step 3, Gibbs sampling is applied to estimate unknown source location. It is to be noted that the obtained rough estimate of source-to-node distance, along with assumed uniformly distributed angular distance are used in generating $\mathbf{G}_{j,k}$. The intuition behind this idea is that once the rough estimate of the source-to-node distance is obtained for a particular node, the source can be anywhere at an angle between $[0, 2\pi]$ with a probability $q_0$.

At the end of the Gibbs sampling, each $j$-th node will generate an estimate to $\hat{\boldsymbol{\theta}}_{j,k}$. When two or more nodes produce same estimate, they are considered to belong to the same *cluster*. In (4.7) and (4.8), large value of $\tau > 0$ implies a large dispersion among the cluster means, whereas parameter $m$ is a prior estimate of these means. Parameter $s$ reflects the uncertainty about the feature measurement errors such that $s$ reflects the confidence in the estimated observation error. The parameter $S$ is the prior estimate of the measurement error.

(4) Step 4 finds out the final estimated source location by discarding outlier cluster if there is any. Location of the winning cluster generates the estimated source location.

A proof of convergence for Algorithm 1 is given in Appendix 4A.

## 4.5 DPMM-based Multiple Diffusive Source Localization

In this section, we will use the best linear unbiased estimator (BLUE) for estimating multiple diffusive source distances from the distributed sensing nodes. The proposed DPMM based Algorithm 1 will be modified and extended to localize multiple static diffusive sources. Let us assume that we have $M$ number of static diffusive sources of same kind with different mass rates in a diffusive field of interest. Therefore, based on the concept developed in section 3.5, substance concentration measurements taken by $j$-th sensor and $k$-th time instant can be written as:

$$\begin{aligned} y_{j,k} &= c_{j,k} + e_{j,k} + b = \sum_{m=1}^{M} c_m(\mathbf{r}_j, t_k) + e_{j,k} + b, \\ &= \sum_{m=1}^{M} \frac{\mu_m}{4\pi\kappa|\mathbf{r}_j - \mathbf{r}_m^s|} \text{erfc}\left(\frac{|\mathbf{r}_j - \mathbf{r}_m^s|}{2\sqrt{\kappa(t_k - t_I)}}\right) + e_{j,k} + b, \end{aligned} \tag{4.9}$$

where erfc(.) is the complementary error function. Since the complementary error function can be approximated as $\text{erfc}(z) \approx 1 - \frac{2}{\sqrt{\pi}}z$, hence our observation model for $j$-th node at the $k$-th time instant can be linearized in terms of the inverse of the distances between the distributed nodes and the diffusive source as:

$$\begin{aligned} y_{j,k} &\approx \sum_{m=1}^{M} \left(\frac{\mu_m}{4\pi\kappa}\right) d_{j,m}^{inv} - \sum_{m=1}^{M} \frac{\mu_m}{4\sqrt{\pi^3\kappa^3(t_k - t_I)}} + b + e_{j,k}, \\ &= \sum_{m=1}^{M} h_m d_{j,m}^{inv} + a_k + e_{j,k}, \end{aligned} \tag{4.10}$$

where $h_m = \frac{\mu_m}{4\pi\kappa}$, $d_{j,m}^{inv} = |\mathbf{r}_j - \mathbf{r}_m^s|^{-1}$ and $a_k = b - \sum_{m=1}^{M} \frac{\mu_m}{4\sqrt{\pi^3\kappa^3(t_k - t_I)}}$. Since all the parameters are known except for the diffusive source locations, we can write $\tilde{y}_{j,k} = y_{j,k} - a_k = \sum_{m=1}^{M} h_m d_{j,m}^{inv} + e_{j,k}$. Therefore, the observation vector formed at the $j$-th node for

$k = 1, 2, \ldots, T$, can be written as:

$$\tilde{\boldsymbol{y}}_j \;=\; \mathbf{H}\mathbf{d}_j^{inv} + \boldsymbol{e}_j, \tag{4.11}$$

where $\mathbf{H} = \frac{1}{4\pi\kappa}\begin{pmatrix} \mu_1 & \mu_2 & \ldots & \mu_M \\ \mu_1 & \mu_2 & \ldots & \mu_M \\ \vdots & \vdots & \ddots & \vdots \\ \mu_1 & \mu_2 & \ldots & \mu_M \end{pmatrix} \in \mathbb{R}^{T\times M}$, $\mathbf{d}_j^{inv} = \left[d_{j,1}^{inv}, d_{j,2}^{inv}, \ldots, d_{j,M}^{inv}\right]^T$, and $\boldsymbol{e}_j = [e_{i1}, e_{i2}, \ldots, e_{iT}]^T$. Since $e_{j,k} \sim \mathcal{N}(0, \sigma^2)$ for $\forall j, k$ and measurement noise is assumed to be independent and identically distributed across space and time, hence the covariance matrix of $\tilde{\boldsymbol{y}}_j$ is $\tilde{\Sigma}_j = \tilde{\Sigma} = \mathrm{diag}\left(\sigma^2, \sigma^2, \ldots, \sigma^2\right) \in \mathbb{R}^{T\times T}$. The optimal BLUE estimator formed at $j$-th node is given by

$$\mathbf{d}_j^{inv} \;=\; \left(\mathbf{H}^T\tilde{\Sigma}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\tilde{\Sigma}^{-1}\tilde{\boldsymbol{y}}_j, \tag{4.12}$$

with estimator variance given by the diagonal elements of $\left(\mathbf{H}^T\tilde{\Sigma}^{-1}\mathbf{H}\right)^{-1}$. Finally, the estimated distances can be obtained by element-wise inverse of the vector $\mathbf{d}_j^{inv}$ as:

$$\hat{\mathbf{d}}_j \;=\; \left[\left(d_{j,1}^{inv}\right)^{-1}, \left(d_{j,2}^{inv}\right)^{-1}, \ldots, \left(d_{j,M}^{inv}\right)^{-1}\right]^T. \tag{4.13}$$

The following algorithm provides a detailed approach for multiple static diffusive source localization using DPMM. Let us denote by $\hat{\boldsymbol{\theta}}_{j,k}^m$, the estimated location of $m$-th source by $j$-th sensor node at $k$-th time instant.

**Algorithm 2** `DPMM-based multiple diffusive source localization`

**STEP 1**: Take measurements $y_{j,k}, \forall j,k$.

**STEP 2**: Obtain source-to-node distance vector $\hat{\mathbf{d}}_j$ for $j \in \{1,2,\ldots,N\}$ using (4.12) and (4.13).

**STEP 3**: For any $k \in \{1,2,\ldots,T\}$

**while** Convergence conditions not satisfied **do**

**for** j= shuffle {1,2,...,N} **do**

For $m = 1,2,\ldots,M$, obtain $\hat{\boldsymbol{\theta}}^m_{j,k}$ using Gibbs sampling, where

$$\hat{\boldsymbol{\theta}}^m_{j,k} \left| \left\{\hat{\boldsymbol{\theta}}^m_{i,k}\right\}_{i\neq j}, \boldsymbol{y}_{1:N} \sim q_0 \mathbf{G}^m_{j,k}(\hat{\boldsymbol{\theta}}^m_{j,k}) + \sum_{i=1,i\neq j}^{N} q_i \delta_{\hat{\boldsymbol{\theta}}^m_{i,k}}, \right.$$

where $\hat{\boldsymbol{\theta}}^m_{j,k} \triangleq \mathbf{r}_{j,k} + \hat{d}^m_{j,k}\angle\psi$ with $\psi \sim \mathcal{U}[0,2\pi]$, $q_0$ and $q_i$'s are defined according to (4.7) and (4.8).

**end for**

**end while**

**STEP 4**: Choose cluster with maximum associated probability for each diffusive source.

A proof of convergence for Algorithm 2 is given in Appendix 4B.

## 4.6 Simulation Results

### 4.6.1 Single Diffusive Source Localization

We consider a 2D diffusive field volume of 2000m$^2$. We assume that the sensors are placed in a uniform 2D grid. Parameters used for simulations are: number of sensing nodes $N$ =64, $\mathbf{r}_0 = [0.01, 0.01]^T$, $\mu$= 100 Kg/s, $b = 0$ Kg/$m^2$, $t_I = 0$ sec and $\kappa = 25m^2/s$. The observation noise is assumed to have Gaussian distribution with mean 0 and variance $\sigma^2 = (0.5 \times 10^{-4})^2$ $Kg^2/m^4$. Number of learning iterations at each time instant is taken to be 60000. Other learning parameters are: $S_j = 1, \forall j$, $s = 3$, $\tau = 1$ and $m = 0$.

The obtained cluster formations from DPMM-based learning process at times $t = 2s$, $t = 4s$, $t = 6s$ and $t = 8s$ are shown in Fig. 4.1 along with their corresponding probabilities and estimated source locations for a single random realization. At each time instant, cluster with the highest probability is considered as the estimated source location. It can be seen that the diffusive source location was successfully estimated in this particular case. In this particular simulation, the static diffusive source can be localized with a resolution of approximately 14 cm.

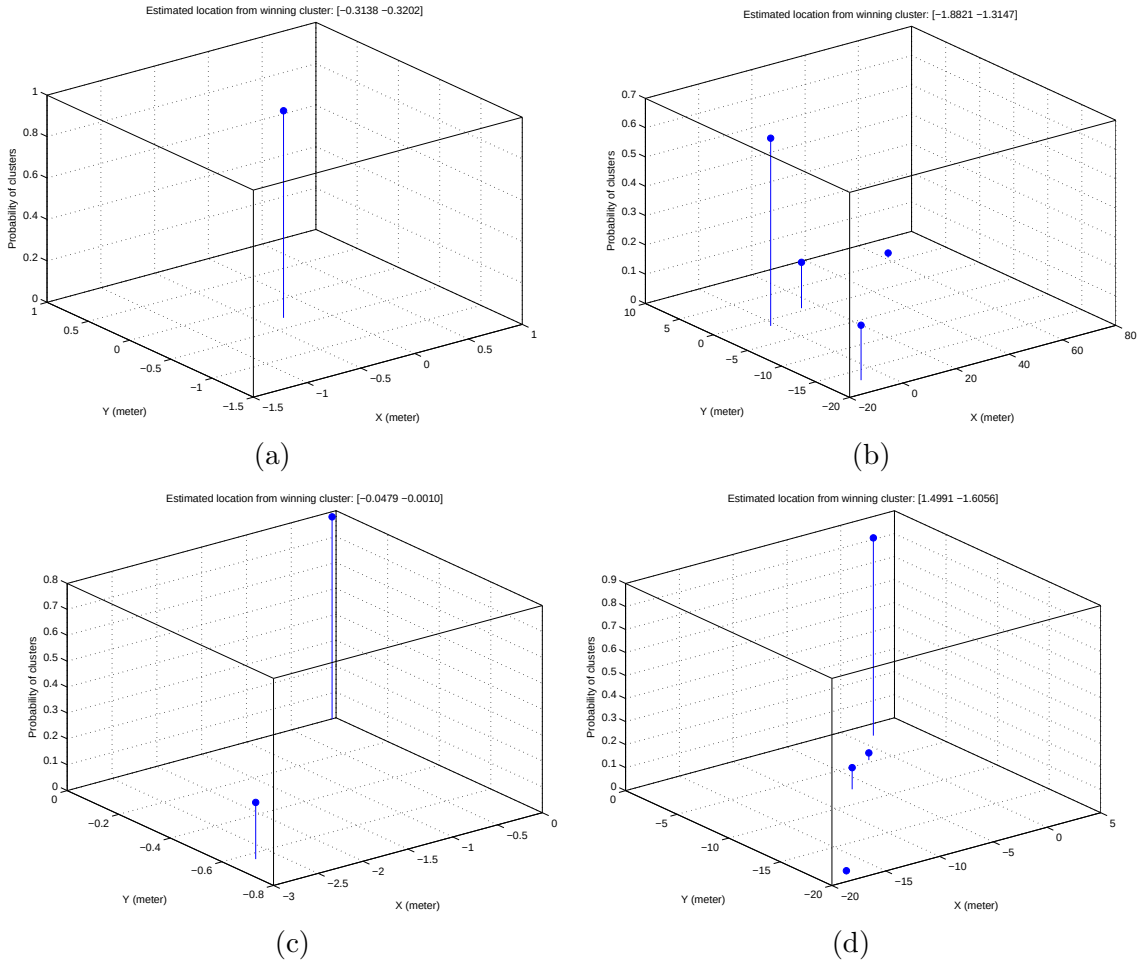



Figure 4.1: Clustering with associated probabilities at time (a) $t = 2$ sec, (b) $t = 4$ sec, (c) $t = 6$ sec, and (d) $t = 8$ sec for a single random realization with actual location being $[0.01, 0.01]^T$.

Figure 4.2 shows performance of the proposed Algorithm 1 in terms of the normalized root-mean-squqred-error (RMSE) as a function of number of sensing nodes, and compares the same with that of the maximum-likelihood estimator (MLE) along with the Cramer-Rao lower bound (CRLB) [40]. The CRLB provides a lower limit on the mean-squared estimation error for non-random parameter estimation for unbiased estimator. In this case, CRLB can be obtained as CRLB $\geq I_{\boldsymbol{\theta}}^{-1}$, where $I_{\boldsymbol{\theta}} \in \mathbb{R}^{2\times 2}$ is the Fisher information matrix (FIM) formed at the FC. The $u-v$th element of the FIM can be found as:

$$[I_{\boldsymbol{\theta}}]_{u,v} = \frac{1}{\sigma^2}\sum_{j=1}^{N}\sum_{k=1}^{T}\left\{\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial \theta_u}\right\}\left\{\frac{\partial c_{j,k}(\boldsymbol{\theta})}{\partial \theta_v}\right\}, \tag{4.14}$$

where (4.14) was obtained assuming independent observations across space and time. Although the performance of Algorithm 1 is satisfactory, it is not as good as its parametric competitor MLE. This is because MLE requires complete description of the likelihood function, which is a very strict requirement. On the other hand, this requirement is relaxed in case of DPMM-based Algorithm 1. In DPMM, knowing the family of noise distribution is useful enough to estimate the parameters related to the distribution. Also the proposed algorithm is computationally less expensive than the MLE. For a particular iteration of any iterative optimization techniques used for the MLE, computational complexity of solving the likelihood function is $O\left(d^2N\right)$, where $d$ is the dimension of the unknown vector to be estimated and $N$ is the sample size. On the other hand, at a particular iteration of the proposed algorithm, the computational complexity of obtaining $\hat{\boldsymbol{\theta}}$ is $O\left(dN\right)$. The computational complexity of of Algorithm 1 is mainly dominated by the 2nd term in (4.6) which requires the evaluation of $N-1$ likelihood functions $q_i$'s.

### 4.6.2 Multiple Diffusive Source Localization

In this case, we assumed that there are two static diffusive sources of same kind present in the diffusive field of interest. Parameters used for simulations are: number of nodes $N$ =64,

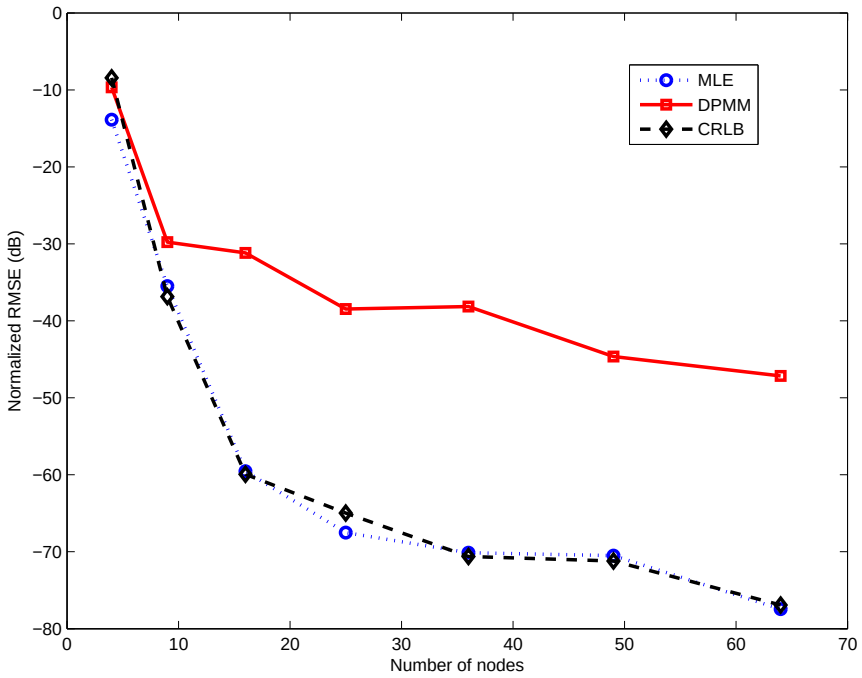
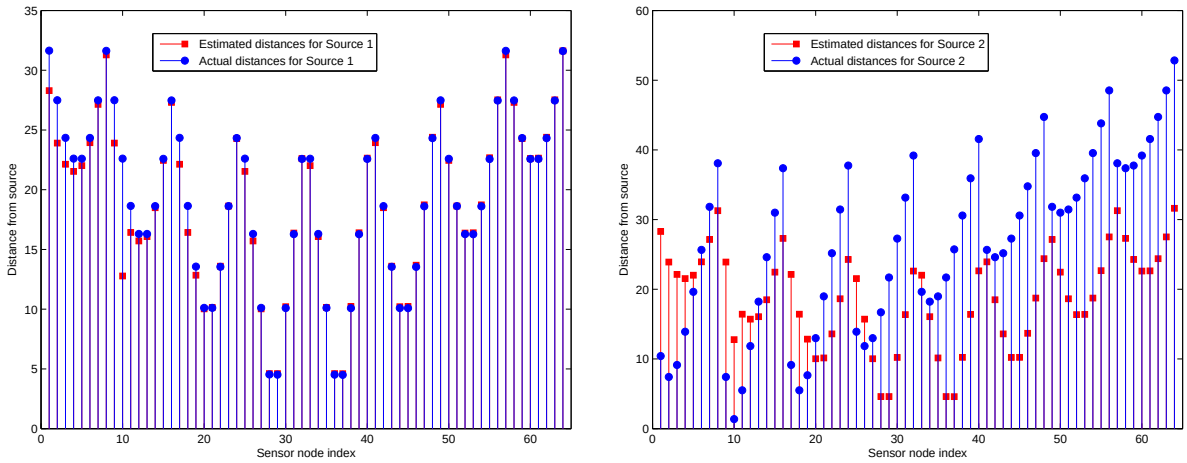



Figure 4.2: Performance of the proposed Algorithm 1.

location of the first source $\mathbf{r}_1^s = [0.01, 0.01]^T$ and the same for the second diffusive source is $\mathbf{r}_2^s = [-20, -20]^T$, mass release rates are $\mu_1$= 1000 Kg/s and $\mu_2$= 100 Kg/s respectively, nuisance parameter $b = 0$, $t_I = 0$ sec and $\kappa = 25m^2/s$. The observation noise is assumed to have Gaussian distribution with mean 0 and variance $\sigma^2 = (1 \times 10^{-5})^2\ Kg^2/m^4$. The measurements are taken at every 1 sec time-step starting from 1 sec and ending at 20 sec. Number of learning iterations at each time instant is taken to be 90000. Other learning parameters are: $S_j = 1, \forall j$, $s = 3$, $\tau = 1$ and $m = 0$.

Figure 4.3 shows the estimated distance using the BLUE estimator based proposed method and the actual distances. It is to be noted that the proposed method for multiple distance estimation is somewhat similar to ICA based source separation problem (*cocktail party problem*), where source mass release rates are considered as signal inputs and the rest of the product term in (4.9) can be considered as mixing terms, which in this case is time-varying. Similar to the ICA-case, the estimated distances from noisy observations are in



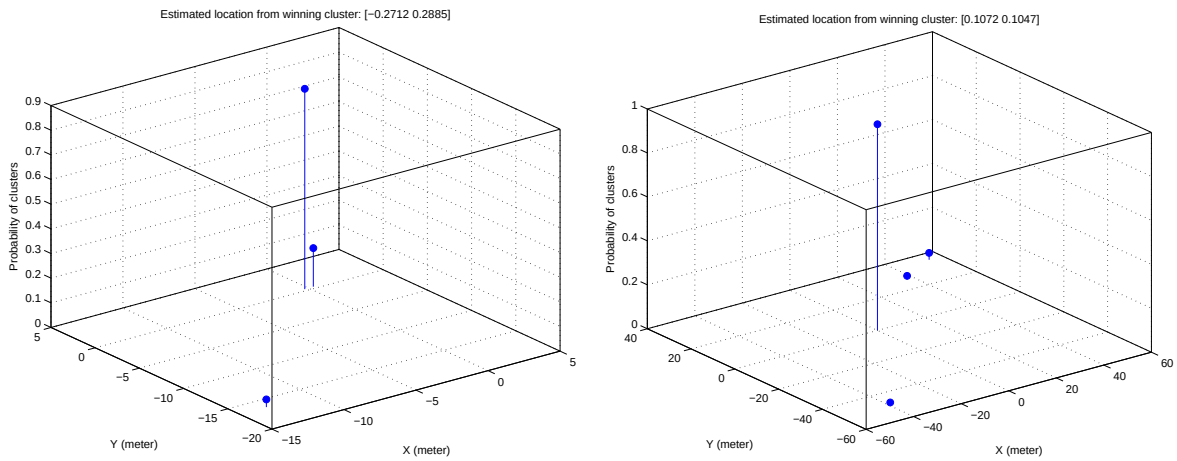

(a) (b)

Figure 4.3: Estimated source-to-node distances for diffusive source (a) One, (b) Two.

fact attenuated version of the actual distances (while the estimated *trend* in the distances are almost the same). As a result post-magnification is needed to obtain a closer solution before the estimated distances are fed into the DPMM-based localization method. It is to be noted that since the maximum distance for our diffusive field of interest can not go beyond $20\sqrt{10}$m, the post-magnification can be done accordingly.

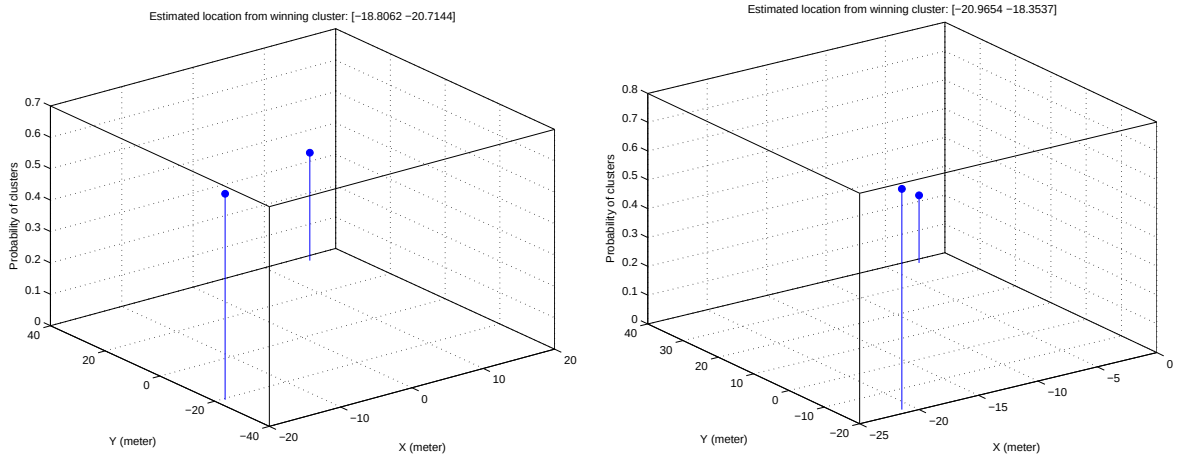
The obtained cluster formations and estimated location of the *first* source from DPMM-based learning process at times $t = 5s$ and $t = 10s$ are given in Figure 4.4 for a single random realization. Cluster formations and estimated location for the *second* source from DPMM-based learning process at times $t = 5s$ and $t = 10s$ are given in Figure 4.5. At each time instant, cluster location with the highest probability is considered as the estimated source location. Figure 4.4 and 4.5 show the clusters formed with their corresponding probabilities and estimated source locations. It can be seen that the diffusive source locations were successfully estimated in each case.

Figure 4.6 shows performance of the proposed algorithm 2 in terms of the total normalized root-mean-squqred-error (RMSE) as a function of number of sensing nodes, and compares




(a) (b)

Figure 4.4: Clustering with associated probabilities at time (a) $t = 5$ sec and (b) $t = 10$ sec for source 1 with actual location being $[0.01, 0.01]^T$ for a single random realization.




(a) (b)

Figure 4.5: Clustering with associated probabilities at time (a) $t = 5$ sec and (b) $t = 10$ sec for source 1 with actual location being $[-20, -20]^T$ for a single random realization.

the same with that of the MLE along with the CRLB. As one would expect, although the performance of Algorithm 2 is satisfactory, it is not as good as its parametric competitor MLE. Also because of the *curse of dimensionality* (increases with the number of sources),

the number of iterations required to reach convergence is higher than that for single diffusive source localization in Algorithm 1. For a particular iteration of any iterative optimization techniques used for the MLE, computational complexity of solving the likelihood function for $M$ number of diffusive sources is $O\left(Md^2NT\right)$, where $d$ is the dimension of the unknown vector to be estimated, $N$ is the sample size and $T$ is the number of time samples. On the other hand, at a particular iteration of the proposed algorithm 2, the computational complexity of obtaining $\hat{\boldsymbol{\theta}}$ is $O\left(Md\max(N,T)\right)$.

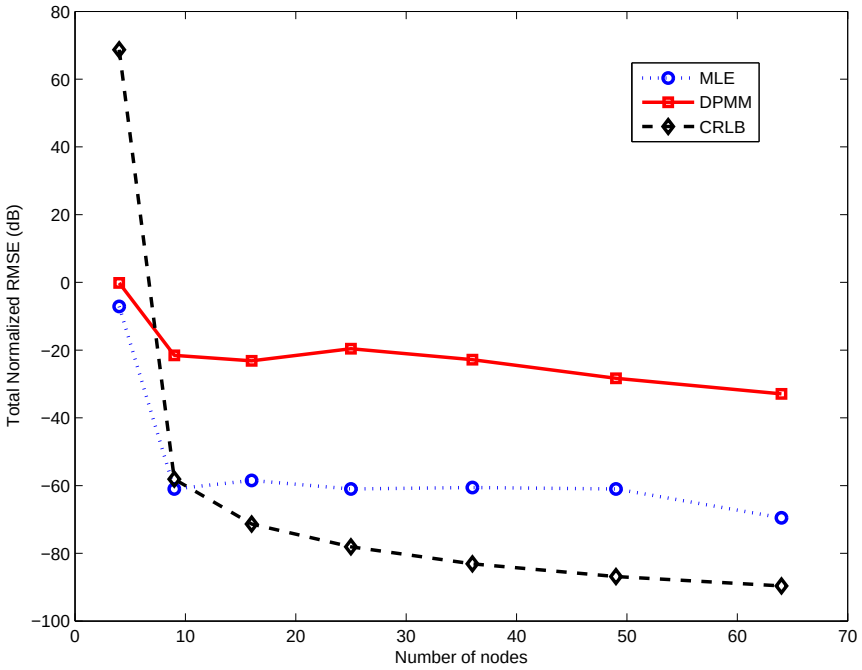



Figure 4.6: Performance of the proposed Algorithm 2.

## 4.7 Conclusion

In this chapter, we proposed two non-parametric diffusive source localization algorithms based on DPMM for single and multiple diffusive sources respectively. We also analytically proved convergence of the proposed algorithms in terms of total variation norm when the

number of iterations in the learning process approach infinity. The performance of the proposed localization methods are shown to be satisfactory using numerical simulations. For the proposed algorithms, knowledge of the family of distribution of the likelihood function is enough. On the other hand, parametric estimation technique would require complete and suitable description of the system model and the likelihood function. It was also found that the proposed algorithms offer low computational complexity for implementation.

# Appendix 4A

## Proof of Convergence for Algorithm 1

To prove the convergence of the proposed Algorithm 1, we will exploit the results given in [77, 79].

Let $Q_I\left(\boldsymbol{\theta}_{1:N}(0), A\right)$ be the probability, with initial value $\boldsymbol{\theta}_{1:N}(0)$, and after one iteration of the algorithm, it produces sample value that is contained in the measurable set $A$, i.e. $Q_I\left(\boldsymbol{\theta}_{1:N}(0), A\right) = P\left\{\boldsymbol{\theta}_{1:N}(1) \in A | \boldsymbol{\theta}_{1:N}(0)\right\}$, where $Q_I(.,.)$ is called the transition kernel of the Markov chain. Also, let $Q_I^u\left(\boldsymbol{\theta}_{1:N}(0), A\right) = P\left\{\boldsymbol{\theta}_{1:N}(u) \in A | \boldsymbol{\theta}_{1:N}(0), u\right\}$ after $u$ number of iterations, and denote the posterior distribution of $\boldsymbol{\theta}_{1:N}$ by $P\left(\boldsymbol{\theta}_{1:N} | \boldsymbol{y}_{1:N}\right)$.

In [77], Theorem 1 states that, for *almost all* starting values of $\boldsymbol{\theta}_{1:N}(0)$, the probability measure $Q_I^u$ defined over the measurable space $\Omega \supset A$, converges in *total variation norm* to the posterior distribution as $u$ goes to infinity. That is, for almost all $\boldsymbol{\theta}_{1:N}(0)$, $\lim_{u\to\infty} \| Q_I^u\left(\boldsymbol{\theta}_{1:N}(0), A\right) - P\left(\boldsymbol{\theta}_{1:N} | \boldsymbol{y}_{1:N}\right) \| = 0$. Since doing the *maximum* operation does not effect the convergence of the algorithm itself, therefore the convergence of the algorithm is ensured. ■

# Appendix 4B

## Proof of Convergence for Algorithm 2

To prove the convergence of the proposed Algorithm 2, we resort to the proof given in Appendix 4A for Algorithm 1.

As convergence was proven for single diffusive source case, likewise in this case, once the estimate of source-to-node distances are obtained using BLUE estimator in step 2 of Algorithm 2, the multiple diffusive source localization problem becomes equivalent to multiple single diffusive source localization in parallel. Therefore, based on Theorem 1 in [77], for *almost all* starting values of $\boldsymbol{\theta}_{1:N}^{m}(0)$, the probability measure $Q_{m,I}^{u}$, $\forall m$, defined over the measurable space $\Omega \supset A$, converges in total variation norm to the posterior distribution $P\left(\boldsymbol{\theta}_{1:N}^{m}|\boldsymbol{y}_{1:N}\right)$ as $u$ goes to infinity. That is, for almost all $\boldsymbol{\theta}_{1:N}^{m}(0)$, $\lim_{u\to\infty} \parallel Q_{m,I}^{u}\left(\boldsymbol{\theta}_{1:N}^{m}(0), A\right) - P\left(\boldsymbol{\theta}_{1:N}^{m}|\boldsymbol{y}_{1:N}\right) \parallel = 0, \forall m$. Again, since taking the *maximum* does not effect the convergence of the algorithm itself, and therefore the convergence of Algorithm 2 is ensured. ■

# Chapter 5

# Moving Diffusive Source Tracking

## 5.1 Introduction

Tracking the state of a moving source/target is an important application of WSN which has garnered tremendous attention in signal processing community. In the Bayesian approach, the key is to construct the posterior probability density function (PDF) of the underlying state vector based on all available information. For linear and Gaussian state dynamics and observation models, the optimal minimum mean squared error (MMSE) solution is tractable and is given by the well-known Kalman filter [40]. However, for most of the real world scenarios, dynamic state estimation problems are nonlinear and non-Gaussian, and obtaining optimal closed-form solution is not tractable under the Bayesian approach. In these cases, suboptimal approached such as extended Kalman filter, Gaussian-sum filter [80] are used with certain approximations. These sub-optimal algorithms become inefficient for highly nonlinear and non-Gaussian systems. In these cases, numerical techniques based on sequential Monte-Carlo methods are used to achieve better performance for highly nonlinear systems. To that end, the idea of particle filtering was introduced in [81] as an effective method of representing PDF in terms of a set of random sampling.

In case of a moving diffusive source tracking, the problem is challenging mainly for two reasons. First, the spatio-temporal concentration distribution is highly nonlinear in terms of the source location. Secondly, the concentration at any time is affected by all past values of the source position. Therefore, time-cumulation effects on the concentrations (i.e. observations) must be taken into account to estimate time-varying parameters. Though use of particle filters for target tracking applications is addressed by many authors in different contexts [35, 37, 38, 41], to the best of our knowledge, moving diffusive source tracking using particle filtering approach has not been attempted before. In this chapter, we propose a particle filter (PF) based target tracking method for moving diffusive source using WSN with a fusion center (FC). The Posterior Cramér-Rao Lower Bound (PCRLB) for the moving source state estimates is also derived as a theoretical performance bound [62].

The remainder of this chapter is organized as follows: Section 5.2 presents the problem formulation and system model.Proposed FC-based moving diffusive source tracking by particle filtering is discussed in detail in section 5.3. Section 5.4 discusses the analytical derivation of the PCRLB as a theoretical performance bound for our proposed tracking scheme. Section 5.5 shows the validity and effectiveness of our proposed methods for diffusive source localization and tracking through numerical simulations. Finally, section 5.6 concludes this chapter by summarizing our results.

## 5.2 Problem Formulation and System Model

### 5.2.1 State Dynamics Model

For the simplicity of exposition and computation, we consider the problem of tracking a diffusive source moving in a 2-dimensional X-Y plane. The assumption can be easily extended to 3-dimensional case without any loss of generality. Let us denote by $\mathbf{s}_k = [x_{s,k}\, y_{s,k}\, \dot{x}_{s,k}\, \dot{y}_{s,k}]^T$, the state vector associated with the moving source at time $t_k$, where the first two elements

represent the source position in 2D and the next two elements represent the speed of the moving source respectively. We assume linear dynamic model for the source state vector:

$$\mathbf{s}_k = \mathbf{F}\mathbf{s}_{k-1} + \mathbf{u}_k, \tag{5.1}$$

for $k = 1, 2, ...$, with the initial known distribution $p(\mathbf{s}_0)$ for $\mathbf{s}_k$, where $\mathbf{F}$ is a $4 \times 4$ matrix that models the state kinematics [82]:

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & T_s & 0 \\ 0 & 1 & 0 & T_s \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{5.2}$$

where $T_s$ is the time difference between two consecutive measurements. The noise vector $\mathbf{u}_k$ is assumed to be zero mean Gaussian with covariance matrix $\mathbf{Q}$ [82]:

$$\mathbf{Q} = \sigma_u^2 \begin{pmatrix} \frac{T_s^3}{3} & 0 & \frac{T_s^2}{2} & 0 \\ 0 & \frac{T_s^3}{3} & 0 & \frac{T_s^2}{2} \\ \frac{T_s^2}{2} & 0 & T_s & 0 \\ 0 & \frac{T_s^2}{2} & 0 & T_s \end{pmatrix}, \tag{5.3}$$

which models the acceleration terms in the spatial directions, and $\sigma_u^2$ is the variance of the process noise.

### 5.2.2 Observation Model

In case of a moving diffusive source continuously emitting diffusing substance in 2D, we may obtain a measurement model for a sensor at a position $\mathbf{r}_{j,k}$ and at time $t_k$ as:

$$\begin{aligned} z_{j,k} &= c(\mathbf{r}_{j,k}, t_k) + \nu(\mathbf{r}_{j,k}, t_k) + b, \text{ for } j \in \mathcal{N} \\ &= c_{j,k} + \nu_{j,k} + b, \end{aligned} \tag{5.4}$$

where $z_{j,k}$ is the $j$-th node's observation at time $t_k$, $c_{j,k} \triangleq c(\mathbf{r}_{j,k}, t_k)$ is the substance concentration at $j$-th node location at time $t_k$ with $c(\mathbf{r}_{j,k}, t_k) = \int_{t_I}^{t_k} \mu(\tau) c_G\left(\mathbf{r}_{j,k} - \mathbf{r}_s(\tau), t - \tau\right)$, moving diffusive source location at time $t_k$ is $\mathbf{r}_{s,k} = \tilde{\mathbf{s}}_k = [x_{s,k}, y_{s,k}]^T$, location of $j$-th node at time $t_k$ is $\mathbf{r}_{j,k} = [x_{j,k}, y_{j,k}]^T$ and $\nu_{j,k} \sim \mathcal{N}(0, \sigma_\nu^2)$ is the sensor measurement noise assumed to be independent in both time and space. Note that for static sensor node locations, we use $\mathbf{r}_{j,k} = \mathbf{r}_j = [x_j, y_j]^T$, by dropping the time index since node locations do not change over time. Assuming additive white Gaussian noise (AWGN) channel for the sake of simplicity, the received signal at the FC from the $j$-th node at time $t_k$ can be written as:

$$\begin{aligned} y_{j,k} &= z_{j,k} + \epsilon_{j,k}, \text{ for } j \in \mathcal{N} \\ &= c_{j,k} + b + \epsilon_{j,k} + \nu_{j,k} = c_{j,k} + b + e_{j,k}, \end{aligned}$$

where $\epsilon_{j,k}$ is the received noise which is assumed to be Gaussian with mean zero, variance $\sigma_\epsilon^2$ and $e_{j,k} = \epsilon_{j,k} + \nu_{j,k}$ and $\sigma^2 = \sigma_\nu^2 + \sigma_\epsilon^2$. We denote $\boldsymbol{y}_{j,1:k}$ as the measurement vector from $j$-th node upto time $t_k$, and $\mathbf{y}_{c,1:k} \triangleq \{\boldsymbol{y}_{1,1:k}, \boldsymbol{y}_{2,1:k}, \ldots, \boldsymbol{y}_{N,1:k}\}^T$ as the collection of all measurements at the FC from $N$ distributed sensor nodes.

In a realistic moving source scenario, the instantaneous velocity is restricted by some practical upper limit. Hence, for lower sampling time $T_s$, we can assume that the moving diffusive source moves in a linear fashion between two observations with an average velocity determined by the source locations $\mathbf{r}_{s,k}$ and $\mathbf{r}_{s,k+1}$. For 2D moving diffusive source tracking with no external force in action, the Green's function can be obtained from (2.10) and (2.11) as:

$$c_G\left(\mathbf{r}_j, t_k\right) = \frac{1}{4\pi\kappa(t_k - t_I)} \exp\left[-\frac{||\mathbf{r}_j - \mathbf{r}_0(t_k)||^2}{4\kappa(t_k - t_I)}\right].$$

Therefore, for a continuous moving diffusive source with constant mass rate $\mu(t) = \mu$, observations taken by the $j$-th node at $k$-th time instant can be written as,

$$y_{j,k} = c_{j,k-1} + \zeta_{j,k} + b + e_{j,k}, \tag{5.5}$$

where

$$\begin{aligned}\zeta_{j,k} &= \mu \int_{t_{k-1}}^{t_k} c_G\left(\mathbf{r}_j - \mathbf{r}_s(\tau), t_k - \tau\right) d\tau, \\ &= \frac{\mu}{4\pi\kappa} \int_{t_{k-1}}^{t_k} \left(\frac{1}{t_k - \tau}\right) \exp\left[-\frac{\left|\left|\mathbf{r}_j - \left\{\mathbf{r}_{s,k-1} + \left(\frac{\mathbf{r}_{s,k} - \mathbf{r}_{s,k-1}}{T_s}\right)(\tau - t_{k-1})\right\}\right|\right|^2}{4\kappa(t_k - \tau)}\right] d\tau.\end{aligned} \tag{5.6}$$

## 5.3 Target Tracking using Particle Filters

In Bayesian belief update, to estimate state vector $\mathbf{s}_k$ at time instant $k$, we need to construct posterior distribution $p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k}\right)$ with initial PDF $p(\mathbf{s}_0)$. The Bayesian belief update is done in two stages: *prediction* and *update*.

*Prediction*: Considering that $p\left(\mathbf{s}_{k-1}|\mathbf{y}_{c,1:k-1}\right)$ is available at time $k$, the PDF $p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k-1}\right)$ can be obtained as [41]:

$$p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k-1}\right) = \int p\left(\mathbf{s}_k|\mathbf{s}_{k-1}\right) p\left(\mathbf{s}_{k-1}|\mathbf{y}_{c,1:k-1}\right) d\mathbf{s}_{k-1}.$$

*Update*: If observations $\mathbf{y}_{c,1:k}$ are available at time instant $k$, the posterior distribution to estimate the state vector $\mathbf{s}_k$ is given by [41]:

$$p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k}\right) = \frac{p\left(\mathbf{y}_{c,k}|\mathbf{s}_k\right) p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k-1}\right)}{p\left(\mathbf{y}_{c,k}|\mathbf{y}_{c,1:k-1}\right)}. \tag{5.7}$$

Since the observation model is highly nonlinear, analytical solution for the optimal estimator is not tractable in our case. Hence, we use sequential Monte Carlo method to approximate the posterior PDF (5.7) with particle filters [81].

Let us denote $\mathcal{X}_k = \{\mathbf{s}_k^i, w_k^i\}_{i=1}^P$ to be the random measure that characterizes the posterior PDF $p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k}\right)$, where $P$ is the number of particles. Then $p\left(\mathbf{s}_k|\mathbf{y}_{c,1:k}\right) \approx \sum_{i=1}^P w_k^i \delta\left(\mathbf{s}_k - \mathbf{s}_k^i\right)$,

where $\delta(.)$ is the Dirac delta function. The state vector estimate at time $t_k$ can be obtained as $\hat{\mathbf{s}}_{k|k} \approx \sum_{i=1}^{P} w_k^i \mathbf{s}_k^i$, and the covariance matrix $U_{k|k}$ of the estimate is $U_{k|k} \approx \sum_{i=1}^{P} w_k^i \left(\mathbf{s}_k^i - \hat{\mathbf{s}}_{k|k}\right)\left(\mathbf{s}_k^i - \hat{\mathbf{s}}_{k|k}\right)^T$. The predicted state $\hat{\mathbf{s}}_{k+1|k}$ and the corresponding covariance matrix $U_{k+1|k}$ can be obtained from the state dynamics in (5.1), as $\hat{\mathbf{s}}_{k+1|k} = \mathbf{F}\hat{\mathbf{s}}_{k|k}$ and $U_{k+1|k} = \mathbf{F}U_{k|k}\mathbf{F}^T + \mathbf{Q}$.

## 5.4 PCRLB Analysis

Analogous to the CRLB, the PCRLB provides a lower bound for the mean-squared error of random parameter estimation [62]. Let us define the joint probability distribution of $\mathbf{S}_k$ and $\mathbf{y}_{c,1:k}$ for an arbitrary $k$ as $p\left(\mathbf{S}_k, \mathbf{y}_{c,1:k}\right) = p_k$, where $\mathbf{y}_{c,1:k}$ is the observation vector formed at the FC at $k$-th time instant and $\mathbf{S}_k = (\mathbf{s}_0, \mathbf{s}_1, \ldots, \mathbf{s}_k)$. Following (5.6), the concentration at any time $k+1$ for any node $j$ can be written as:

$$c(\mathbf{r}_j, t_{k+1}) \triangleq c_{j,k+1} = \zeta_{j,0:1} + \zeta_{j,1:2} + \ldots + \zeta_{j,k-1:k} + \zeta_{j,k:k+1}.$$

Based on the assumed observation model in (5.5), the log-likelihood function $\mathcal{L}_{k+1}$ at $(k+1)$-th time instant formed at the FC is given by

$$\begin{aligned}\mathcal{L}_{k+1} &= \log p\left(\mathbf{y}_{c,k+1} | \mathbf{s}_{k+1}, \mathbf{S}_k\right), \\ &= -\frac{N}{2}\log(2\pi\sigma^2) - \sum_{j=1}^{N} \frac{1}{2\sigma^2}\left(y_{j,k+1} - c\left(\mathbf{r}_j, t_k\right) - \zeta_{j,k+1} - b\right)^2.\end{aligned}$$

Let $I(\mathbf{S}_k) \in \mathbf{R}^{4k \times 4k}$ be the information matrix derived from the joint distribution $p_k$. We wish to solve for the information submatrix for estimating $\mathbf{s}_k$, denoted by $I_k$. The following theorem gives a two-step recipe for computing $I_k$.

*Theorem 3*: The sequence $\{I_{k+1}\}$ of the posterior information submatrices for estimating

state vectors $\mathbf{s}_{k+1}$ can be computed as follows:

$$I_{k+1} = D_{k+1} - \begin{bmatrix} L_{k+1} & -\mathbf{Q}^{-1}\mathbf{F} + M_{k+1} \end{bmatrix} [I(\mathbf{S}_k) + R_{k+1}]^{-1} \begin{bmatrix} L_{k+1}^T \\ -\mathbf{F}^T\mathbf{Q}^{-1} + M_{k+1}^T \end{bmatrix} \quad (5.8)$$

where $M_{k+1} = -\mathbb{E}\left\{\Delta^{\mathbf{s}_k}_{\mathbf{s}_{k+1}} \mathcal{L}_{k+1}\right\}$, $D_{k+1} = -\mathbb{E}\left\{\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_{k+1}} \log p_{k+1}\right\}$,
$L_{k+1} = \left[-\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_{k+1}} \mathcal{L}_{k+1}\right\} - \mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_{k+1}} \mathcal{L}_{k+1}\right\} \ldots - \mathbb{E}\left\{\Delta^{\mathbf{s}_{k-1}}_{\mathbf{s}_{k+1}} \mathcal{L}_{k+1}\right\}\right]$,

$$R_{k+1} = \begin{bmatrix} -\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_0} \mathcal{L}_{k+1}\right\} & -\mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_0} \mathcal{L}_{k+1}\right\} & \ldots & -\mathbb{E}\left\{\Delta^{\mathbf{s}_k}_{\mathbf{s}_0} \mathcal{L}_{k+1}\right\} \\ -\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_1} \mathcal{L}_{k+1}\right\} & -\mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_1} \mathcal{L}_{k+1}\right\} & \ldots & -\mathbb{E}\left\{\Delta^{\mathbf{s}_k}_{\mathbf{s}_1} \mathcal{L}_{k+1}\right\} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_k} \mathcal{L}_{k+1}\right\} & -\mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_k} \mathcal{L}_{k+1}\right\} & \ldots & -\mathbb{E}\left\{\Delta^{\mathbf{s}_k}_{\mathbf{s}_k} \mathcal{L}_{k+1}\right\} + \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} \end{bmatrix}, \quad (5.9)$$

and $\Delta^{\Theta}_{\Phi} = \nabla_{\Phi}\nabla^T_{\Theta}$ with $\nabla$ being the Laplacian operator.

*Proof*: See Appendix 5A.

Note that the information submatrix computation in (5.8) involves computation of the inverse of a matrix of size $4k \times 4k$. This is because of the output $y_{j,k+1}$ at the $j$-th node at $(k+1)$-th time instant being a function of all the previous states $\mathbf{S}_{k+1}$.

## 5.5 Simulation Results

In this section, we analyze the performance of our proposed moving diffusive source tracking scheme. For the sake of simplicity, we consider a 2D diffusive field volume of $\Lambda = [-50, 50] \times [-50, 50]$ m$^2$. We assume that the sensors are placed in a uniform 2D grid such that the distance between adjacent sensors along the same ordinate is approximately 14.3 m. Parameters used for simulations are: number of nodes $N$ =64, $\mathbf{r}_0 = [0, 0]^T$, $\mu$= 1000 Kg/s, $b = 10^{-4}$ Kg/$m^2$, $t_I = 0$ sec and $\kappa = 25m^2/s$. The observation noise is assumed to have

Gaussian distribution with mean 0 and variance $\sigma^2 = 1 \times 10^{-4}\ Kg^2/m^4$. The measurements are taken at every 0.5 sec time-step starting from 0.5 sec and ending at 30 sec.

The initial source state vector is assumed to be Gaussian with mean $\mu = [0, 0, 0, 0]^T$ and covariance matrix $\Sigma_0 = diag\left([0.01, 0.01, 0.01, 0.01]^T\right)$. The intensity of the state process noise is $\sigma_u^2 = 0.1$. Sampling time is assumed to be $T_s = 0.5$sec. Total number of random realizations used for simulations is 50. The tracking is performed for 30 sec and the number of particles in the particle filter (PF) is $N_p = 1000$. The performance measure is taken as the root-mean-squared-error (RMSE) of the moving source position estimate given by $RMSE_k = \sqrt{(x_{s,k} - \hat{x}_{s,k})^2 + (y_{s,k} - \hat{y}_{s,k})^2}$. The RMSE is compared with the square root of the PCRLB components of the position error, $PCRLB_k \approx \sqrt{\left[I_k^{-1}\right]_{11} + \left[I_k^{-1}\right]_{22}}$.

Figure 5.1 and 5.2 show the tracking performances of the proposed tracking scheme using particle filter for grid-based and random node deployment strategies respectively. It can be seen that the target trajectory can be tracked with better accuracy in Figure 5.2 with compared to that in Figure 5.1. Figure 5.2(b) and 5.1(b) show the RMSE's on the tracking performances for the aforementioned two node deployment strategies respectively. The obtained RMSE with the random node deployment case is better and closer to the derived PCRLB than those for the grid-based node deployment case. This is because for a fixed node density, the expected nearest neighbor node distance (from the source) in case of random node deployment is less than the inter-node spacing in grid-based node deployment, which in our case is 14.3 m. The random node deployment is specially suitable when there is no pre-designed infrastructure for sensor network and also when the diffusive field is hazardous for human deployment.

It is of interest also to investigate the performance of the proposed target tracking method when the sampling time $T_s$ is varying. Figure 5.3 shows the effect of sampling time $T_s$ on the tracking performances of the proposed moving diffusive source tracking scheme using grid-based node deployment strategy keeping all the other parameters same as mentioned before. As one would expect, the tracking performance decrease with the increase of sampling

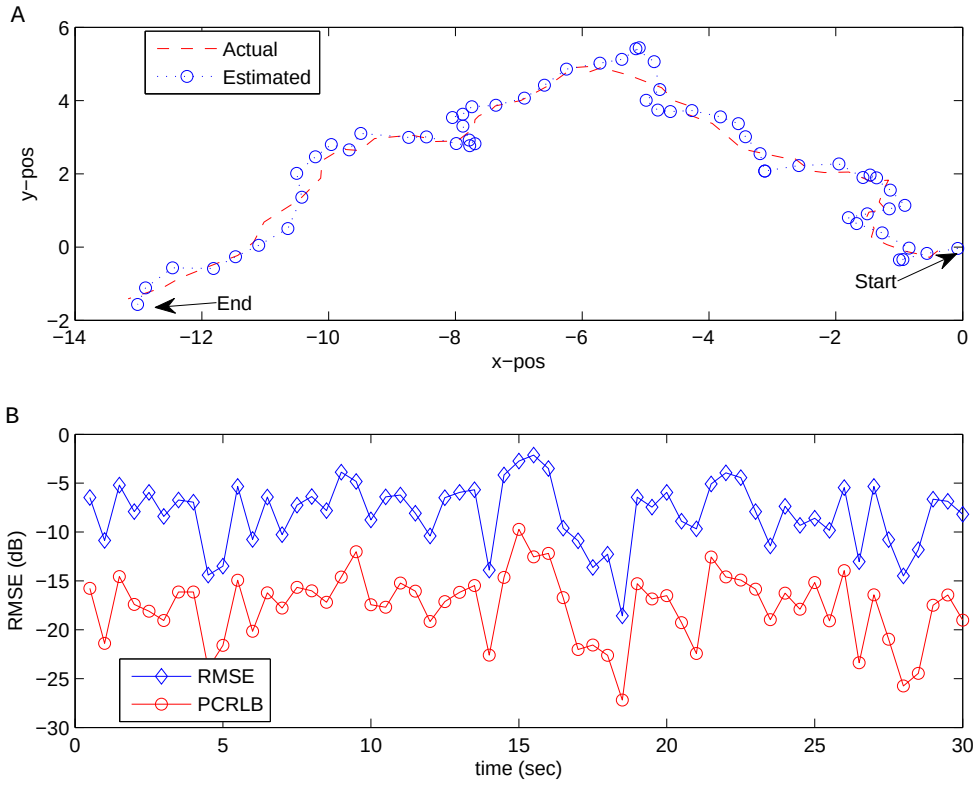



Figure 5.1: (a) Actual and estimated trajectories of the moving diffusive source, and (b) RMSE (dB) for grid-based sensor node deployment.

time $T_s$. This is because for higher values of $T_s$, the process noise will increase according to (5.3). Since we are also assuming that the movement of the diffusive source is almost linear between two successive time instant, the lower $T_s$ will result in better accuracy of the proposed tracking scheme.

## 5.6 Conclusion

In this chapter, we proposed a particle filter based target tracking method for moving diffusive source emitting substance continuously into a dispersive medium. The PCRLB correspond-

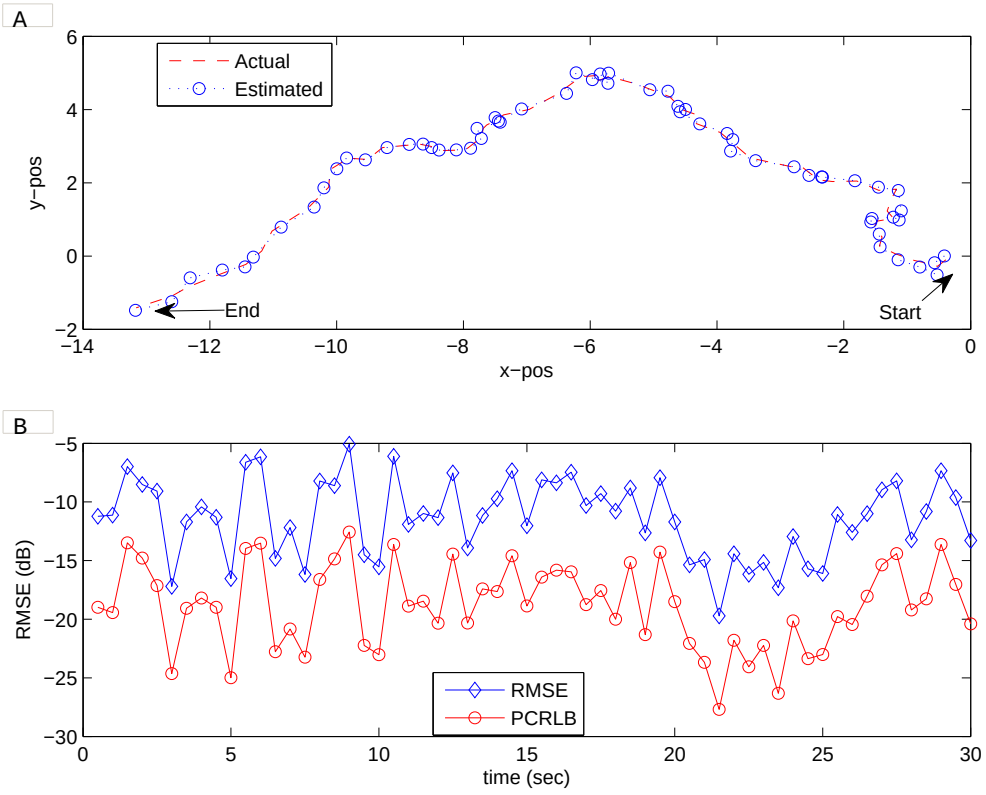



Figure 5.2: (a) Actual and estimated trajectories of the moving diffusive source, and (b) RMSE (dB) for random sensor node deployment.

ing to moving diffusive source tracking was obtained as a theoretical performance measure and compared with the simulation results. Both grid-based and random node deployment strategies were investigated for our proposed tracking scheme. The effect of sampling time on the moving source tracking was also studied. The performance of the proposed estimation and tracking methods are shown to be excellent using numerical simulations.

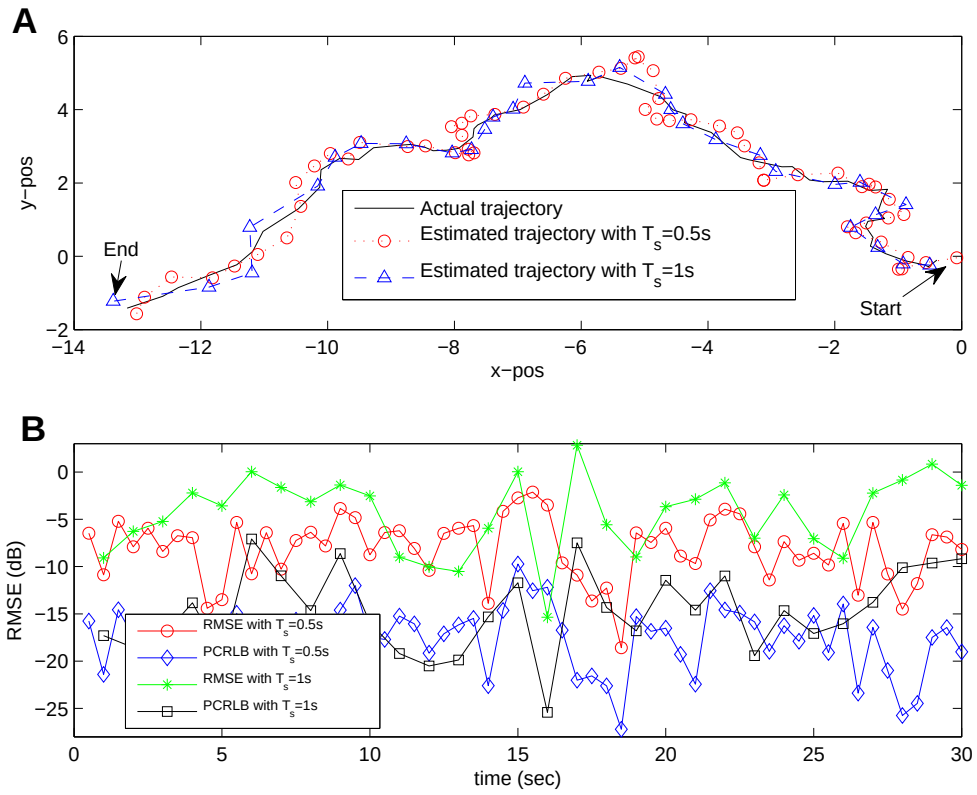



Figure 5.3: (a) Actual and estimated trajectories of the moving diffusive source, and (b) RMSE (dB) for different values of sampling time $T_s$.

# Appendix 5A

## Proof of Theorem 3

The joint probability distribution of $\mathbf{S}_k$ and $\mathbf{y}_{c,1:k}$ at $k$-th time instant can be written as:

$$p\left(\mathbf{S}_k, \mathbf{y}_{c,1:k}\right) \;=\; p_k = p(\mathbf{s}_0) \prod_{m=1}^{k} p\left(\mathbf{y}_{c,1:k} \mid \mathbf{s}_{1:m}\right) \prod_{n=1}^{k} p\left(\mathbf{s}_n \mid \mathbf{s}_{n-1}\right).$$

For $p(\mathbf{s}_0) \sim \mathcal{N}(\mu_0, \mathbf{\Sigma}_0)$, the initial condition for the FIM is $I(\mathbf{S}_0) = \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_0} \log p(\mathbf{s}_0)\right\} = \mathbf{\Sigma}_0^{-1}$. Decomposing $\mathbf{S}_1$ as $\mathbf{S}_1 = [\mathbf{s}_0^T, \mathbf{s}_1^T]^T$, $I(\mathbf{S}_1)$ can be obtained as

$$\begin{aligned} I(\mathbf{S}_1) &= \begin{bmatrix} \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_0} \log p_1\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_1} \log p_1\right\} \\ \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_0} \log p_1\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_1} \log p_1\right\} \end{bmatrix}, \\ &= \begin{bmatrix} I(\mathbf{S}_0) + R_1 & \left[\mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_1} \mathcal{L}_1\right\} - \mathbf{Q}^{-1}\mathbf{F}\right]^T \\ \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_1} \mathcal{L}_1\right\} - \mathbf{Q}^{-1}\mathbf{F} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_1} \mathcal{L}_1\right\} + \mathbf{Q}^{-1} \end{bmatrix} \triangleq \begin{bmatrix} A_1 & B_1^T \\ B_1 & D_1 \end{bmatrix}. \end{aligned}$$

Since error is independent across space and time, using concept from block matrix inversion, the information submatrix that provides the mean square error estimate of $\mathbf{s}_1$ is given by

$$I_1 = D_1 - B_1 A_1^{-1} B_1^T = D_1 - B_1 \left[I(\mathbf{S}_0) + R_1\right]^{-1} B_1^T,$$

where $D_1 = \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_1} \mathcal{L}_1\right\} + \mathbf{Q}^{-1}$, $R_1 = \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_0} \mathcal{L}_1\right\} + \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F}$ and $B_1 = \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_0} \mathcal{L}_1\right\} - \mathbf{F}^T\mathbf{Q}^{-1}$. Similarly, decomposing $\mathbf{S}_2$ as $\mathbf{S}_2 = [\mathbf{s}_0^T, \mathbf{s}_1^T, \mathbf{s}_2^T]^T$, the FIM $I(\mathbf{S}_2)$ can be written as follows:

$$\begin{aligned} I(\mathbf{S}_2) &= \begin{bmatrix} \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_0} \log p_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_1} \log p_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_2} \log p_2\right\} \\ \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_0} \log p_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_1} \log p_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_2} \log p_2\right\} \\ \mathbb{E}\left\{-\Delta_{\mathbf{s}_2}^{\mathbf{s}_0} \log p_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_2}^{\mathbf{s}_1} \log p_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_2}^{\mathbf{s}_2} \log p_2\right\} \end{bmatrix}, \\ &= \begin{bmatrix} A_1 + \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_0} \mathcal{L}_2\right\} & B_1^T + \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_1} \mathcal{L}_2\right\} & \mathbb{E}\left\{-\Delta_{\mathbf{s}_0}^{\mathbf{s}_2} \mathcal{L}_2\right\} \\ B_1 + \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_0} \mathcal{L}_2\right\} & D_1 + \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} + \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_1} \mathcal{L}_2\right\} & -\mathbf{F}^T\mathbf{Q}^{-1} + \mathbb{E}\left\{-\Delta_{\mathbf{s}_1}^{\mathbf{s}_2} \mathcal{L}_2\right\} \\ \mathbb{E}\left\{-\Delta_{\mathbf{s}_2}^{\mathbf{s}_0} \mathcal{L}_2\right\} & -\mathbf{Q}^{-1}\mathbf{F} + \mathbb{E}\left\{-\Delta_{\mathbf{s}_2}^{\mathbf{s}_1} \mathcal{L}_2\right\} & \mathbf{Q}^{-1} + \mathbb{E}\left\{-\Delta_{\mathbf{s}_2}^{\mathbf{s}_2} \mathcal{L}_2\right\} \end{bmatrix}, \\ &\triangleq \begin{bmatrix} I(\mathbf{S}_1) + R_2 & B_2^T \\ B_2 & D_2 \end{bmatrix}. \end{aligned} \tag{5.10}$$

The information submatrix $I_2$ can be found as an inverse of the right-lower $4 \times 4$ submatrix of $[I(\mathbf{S}_2)]^{-1}$:

$$I_2 = D_2 - B_2 \left[I(\mathbf{S}_1) + R_2\right]^{-1} B_2^T,$$

where $D_2 = \mathbf{Q}^{-1} + \mathbb{E}\left\{-\Delta^{\mathbf{s}_2}_{\mathbf{s}_2}\mathcal{L}_2\right\}$, $B_2 = \left[\mathbb{E}\left\{-\Delta^{\mathbf{s}_0}_{\mathbf{s}_2}\mathcal{L}_2\right\} - \mathbf{Q}^{-1}\mathbf{F} + \mathbb{E}\left\{-\Delta^{\mathbf{s}_1}_{\mathbf{s}_2}\mathcal{L}_2\right\}\right]$, and

$$R_2 = \begin{bmatrix} \mathbb{E}\left\{-\Delta^{\mathbf{s}_0}_{\mathbf{s}_0}\mathcal{L}_2\right\} & \mathbb{E}\left\{-\Delta^{\mathbf{s}_1}_{\mathbf{s}_0}\mathcal{L}_2\right\} \\ \mathbb{E}\left\{-\Delta^{\mathbf{s}_0}_{\mathbf{s}_1}\mathcal{L}_2\right\} & \mathbf{F}^T\mathbf{Q}^{-1}\mathbf{F} + \mathbb{E}\left\{-\Delta^{\mathbf{s}_1}_{\mathbf{s}_1}\mathcal{L}_2\right\} \end{bmatrix}.$$

By extending the above procedure and decomposing $\mathbf{S}_{k+1} = \left[\mathbf{s}_0^T, \mathbf{s}_1^T, \ldots, \mathbf{s}_{k+1}^T\right]^T$, $I\left(\mathbf{S}_{k+1}\right)$ can be obtained as:

$$\begin{aligned} I(\mathbf{S}_{k+1}) &= \begin{bmatrix} -\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_0}\log p_{k+1}\right\} & -\mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_0}\log p_{k+1}\right\} & \ldots & -\mathbb{E}\left\{\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_0}\log p_{k+1}\right\} \\ -\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_1}\log p_{k+1}\right\} & -\mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_1}\log p_{k+1}\right\} & \ldots & -\mathbb{E}\left\{\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_1}\log p_{k+1}\right\} \\ \vdots & \vdots & \ddots & \vdots \\ -\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_{k+1}}\log p_{k+1}\right\} & -\mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_{k+1}}\log p_{k+1}\right\} & \ldots & -\mathbb{E}\left\{\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_{k+1}}\log p_{k+1}\right\} \end{bmatrix}, \\ &\triangleq \begin{bmatrix} I(\mathbf{S}_k) + R_{k+1} & \begin{bmatrix} L_{k+1}^T \\ -\mathbf{F}^T\mathbf{Q}^{-1} + M_{k+1}^T \end{bmatrix} \\ \begin{bmatrix} L_{k+1} & -\mathbf{Q}^{-1}\mathbf{F} + M_{k+1} \end{bmatrix} & D_{k+1} \end{bmatrix}. \end{aligned} \tag{5.11}$$

The information submatrix $I_{k+1}$ can be generalized as an inverse of the right-lower $4 \times 4$ submatrix of $[I(\mathbf{S}_{k+1})]^{-1}$ in (5.11), where $M_{k+1} = -\mathbb{E}\left\{\Delta^{\mathbf{s}_k}_{\mathbf{s}_{k+1}}\mathcal{L}_{k+1}\right\}$,
$D_{k+1} = -\mathbb{E}\left\{\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_{k+1}}\log p_{k+1}\right\} = \mathbf{Q}^{-1} + \mathbb{E}\left\{-\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_{k+1}}\mathcal{L}_{k+1}\right\} \triangleq \mathbf{Q}^{-1} + \tilde{D}_{k+1}$,
$L_{k+1} = \left[-\mathbb{E}\left\{\Delta^{\mathbf{s}_0}_{\mathbf{s}_{k+1}}\mathcal{L}_{k+1}\right\} - \mathbb{E}\left\{\Delta^{\mathbf{s}_1}_{\mathbf{s}_{k+1}}\mathcal{L}_{k+1}\right\} \ldots - \mathbb{E}\left\{\Delta^{\mathbf{s}_{k-1}}_{\mathbf{s}_{k+1}}\mathcal{L}_{k+1}\right\}\right]$, and $R_{k+1}$ is defined in (5.9). The only non-zero elements of $\tilde{D}_{k+1} = \mathbb{E}\left\{-\Delta^{\mathbf{s}_{k+1}}_{\mathbf{s}_{k+1}}\mathcal{L}_{k+1}\right\} \in \mathbb{R}^{4\times 4}$ are given by,

$$\begin{aligned} \left[\tilde{D}_{k+1}\right]_{11} &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial\zeta_{j,k:k+1}}{\partial\mathbf{s}_{k+1}(1)}\right]^2, \\ \left[\tilde{D}_{k+1}\right]_{12} &= \left[\tilde{D}_{k+1}\right]_{21} = \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial\zeta_{j,k:k+1}}{\partial\mathbf{s}_{k+1}(1)}\right]\left[\frac{\partial\zeta_{j,k:k+1}}{\partial\mathbf{s}_{k+1}(2)}\right], \\ \left[\tilde{D}_{k+1}\right]_{22} &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial\zeta_{j,k:k+1}}{\partial\mathbf{s}_{k+1}(2)}\right]^2. \end{aligned}$$

Similarly the only non-zero elements of $M_{k+1} = -\mathbb{E}\left\{\Delta^{\mathbf{s}_k}_{\mathbf{s}_{k+1}} \mathcal{L}_{k+1}\right\} \in \mathbb{R}^{4\times 4}$, can be obtained as

$$\begin{aligned}
[M_{k+1}]_{11} &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k+1}(1)}\right]\left[\frac{\partial \zeta_{j,k-1:k}}{\partial \mathbf{s}_{k}(1)} + \frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k}(1)}\right], \\
[M_{k+1}]_{12} &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k+1}(1)}\right]\left[\frac{\partial \zeta_{j,k-1:k}}{\partial \mathbf{s}_{k}(2)} + \frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k}(2)}\right], \\
[M_{k+1}]_{21} &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k+1}(2)}\right]\left[\frac{\partial \zeta_{j,k-1:k}}{\partial \mathbf{s}_{k}(1)} + \frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k}(1)}\right], \\
[M_{k+1}]_{22} &= \frac{1}{\sigma^2}\sum_{j=1}^{N}\left[\frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k+1}(2)}\right]\left[\frac{\partial \zeta_{j,k-1:k}}{\partial \mathbf{s}_{k}(2)} + \frac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k}(2)}\right],
\end{aligned}$$

where the partial-derivative components are defined as follows using (5.6):

$$\left.\begin{aligned}
\tfrac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_k(1)} &= \tfrac{\mu}{8\pi T_s \kappa^2}\int_{t_k}^{t_{k+1}} \tfrac{|x_j - x_0(\tau)|}{(t_{k+1}-\tau)^2}\exp\left[-\tfrac{\left|\mathbf{r}_j - \left\{\mathbf{r}_s(t_k) + \left(\frac{\mathbf{r}_s(t_{k+1})-\mathbf{r}_s(t_k)}{T_s}\right)(\tau - t_k)\right\}\right|^2}{4\kappa(t_{k+1}-\tau)}\right](t_{k+1}-\tau)d\tau, \\
\tfrac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_k(2)} &= \tfrac{\mu}{8\pi T_s \kappa^2}\int_{t_k}^{t_{k+1}} \tfrac{|y_j - y_s(\tau)|}{(t_{k+1}-\tau)^2}\exp\left[-\tfrac{\left|\mathbf{r}_j - \left\{\mathbf{r}_s(t_k) + \left(\frac{\mathbf{r}_s(t_{k+1})-\mathbf{r}_s(t_k)}{T_s}\right)(\tau - t_k)\right\}\right|^2}{4\kappa(t_{k+1}-\tau)}\right](t_{k+1}-\tau)d\tau, \\
\tfrac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k+1}(1)} &= \tfrac{\mu}{8\pi T_s \kappa^2}\int_{t_k}^{t_{k+1}} \tfrac{|x_j - x_s(\tau)|}{(t_{k+1}-\tau)^2}\exp\left[-\tfrac{\left|\mathbf{r}_j - \left\{\mathbf{r}_s(t_k) + \left(\frac{\mathbf{r}_s(t_{k+1})-\mathbf{r}_s(t_k)}{T_s}\right)(\tau - t_k)\right\}\right|^2}{4\kappa(t_{k+1}-\tau)}\right](\tau - t_k)d\tau, \\
\tfrac{\partial \zeta_{j,k:k+1}}{\partial \mathbf{s}_{k+1}(2)} &= \tfrac{\mu}{8\pi T_s \kappa^2}\int_{t_k}^{t_{k+1}} \tfrac{|y_j - y_s(\tau)|}{(t_{k+1}-\tau)^2}\exp\left[-\tfrac{\left|\mathbf{r}_j - \left\{\mathbf{r}_s(t_k) + \left(\frac{\mathbf{r}_s(t_{k+1})-\mathbf{r}_s(t_k)}{T_s}\right)(\tau - t_k)\right\}\right|^2}{4\kappa(t_{k+1}-\tau)}\right](\tau - t_k)d\tau,
\end{aligned}\right\} \quad (5.12)$$

and

$$\begin{aligned}
x_s(\tau) &= \left(\frac{t_{k+1}-\tau}{T_s}\right)x_s(t_k) + \left(\frac{\tau - t_k}{T_s}\right)x_s(t_{k+1}), \\
y_s(\tau) &= \left(\frac{t_{k+1}-\tau}{T_s}\right)y_s(t_k) + \left(\frac{\tau - t_k}{T_s}\right)y_s(t_{k+1}).
\end{aligned}$$

Following the same approach as above, the elements of the matrix $L_{k+1} \in \mathbb{R}^{4\times 4k}$ can easily be obtained at each time instant. ■

# Chapter 6

# Game Theoretic Lifetime Improvement of WSN in Estimation

## 6.1 Introduction

Wireless Sensor networks tend to experience premature failure since some nodes might run out of their batteries rapidly due to work load variations, communication environments or hardware setup. It is undesirable for a sensor node to waste power as excessive use of battery power can shorten the lifetime of a node. Resource-constrained WSNs rely on collaborative signal and information processing for efficient handling of large volumes of data collected by distributed sensor nodes. Node collaboration, however, requires inter sensor communication. Payoff and cost of collaboration can be modeled, respectively, as the improved quality of processed outputs and the required power or bandwidth for communication. Thus there needs to be a trade-off between performance and cost of collaborative information processing.

The rich collection of tools from cooperative game theory can be very useful in approaching such collaborative signal processing problems in a sensor network. Unlike non-cooperative game theory, where individual players compete with each other to achieve their goals of max-

imizing individual payoffs, cooperative game theory allows competing players (or nodes) to form coalitions so as to efficiently achieve their individual goals. For a resource constrained WSN, cooperative game theory can be a natural choice and comes in handy when estimating a parameter with desired estimator quality is the ultimate goal. In current literature, very few attempts have been made to exploit the rich collection of cooperative game theory in power/energy-constrained WSNs tasked with estimating a parameter. For example, in [56], a novel concept of incompletely cooperative game theory was used to simultaneously achieve energy conservation and throughput for WSNs. On the other hand, [57,58] used cooperative game theory for channel/bandwidth allocation problem. In this chapter, we use tools from cooperative game theory to develop a formal analytical framework for fair allocation of power among participating sensor nodes to achieve a sequential estimation task while at the same time maximizing overall network lifetime. In particular, we use the concept of the Shapley value [61] to achieve power allocation among distributed nodes with power constraints.

The remainder of this chapter is organized as follows. In section 6.2, we present the sensor network model for the sequential estimation problem. Section 6.3 discusses about the basic concepts and theorems for the Shapley value based solution method. The proposed power allocation algorithm with a combined objectives of network lifetime improvement and estimation, is discussed in section 6.4. Section 6.5 evaluates the performance of the proposed solution via simulations. Finally, Section 6.6 concludes the chapter by summarizing our results.

## 6.2 Sensor Network Model for Sequential Estimation

We consider a sensor network consisting of a Fusion Center (FC) and $N$ nodes tasked with estimating a non-random parameter $\theta$ sequentially as shown in Fig. 6.1. We consider the FC itself as a node with its own estimate of the parameter $\theta$ and is denoted as node-0. The set of distributed nodes are denoted as $\mathcal{N} = \{1, 2, \cdots, N\}$. Objective of each node, or a

set of nodes, is to obtain a reliable estimation of $\theta$. Sensor network may consists of mobile nodes or a hybrid of fixed and mobile nodes and the wireless channel can be time-varying. However, only a *quasi-static network* is considered here, in which node locations as well as fading coefficients can be assumed fixed for a certain period of time, whereas from block to block they could be varying.

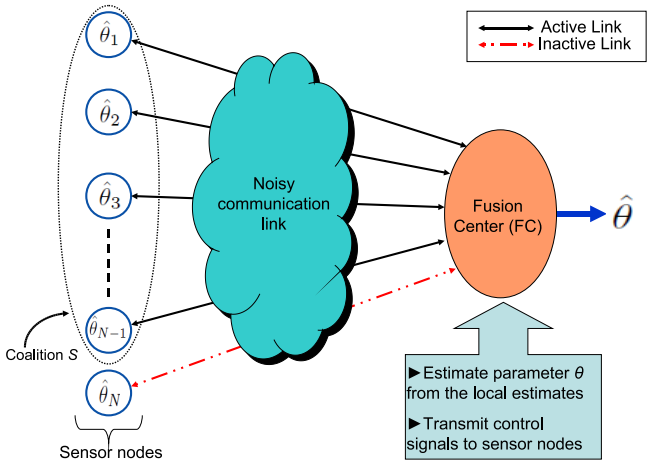



Figure 6.1: A typical WSN architecture with a FC.

Local estimator at node $i$ is denoted by $\hat{\theta}_i$, for $i = 0, 1, 2, \cdots, N$. All local estimators are assumed to be unbiased and their respective variances denoted by $V_i$. Under the quasi-static assumption, we may assume that at the beginning of each block, the FC (node-0) has access to the quality of estimates at the distributed nodes as given by $V_i$'s for $i \in \mathcal{N}$. It is assumed that the FC forms its updated estimator sequentially (in a predetermined order or randomly) by combining its own observation with the noise-corrupted estimates of the selected distributed nodes received over noisy communication links. The FC keeps on sequentially updating its estimator until it achieves a certain predetermined estimation quality denoted by $V_t$. We consider the case when the distributed nodes are not supposed to transmit at their maximum powers, while the goal is to achieve a desired quality of estimate at the FC using as fewer a number of nodes as possible, and at faster a rate as possible. The

objective is to define a fair allocation of network resources for collaborating nodes in terms of their transmit powers. The sensor nodes in the assumed WSN are powered by batteries with limited lifetime, which is dissipated during the data transmission/reception. Note that, we define the lifetime of a sensor network as the time after which at least one or a certain fraction of sensor nodes run out of their batteries, resulting in a hole within the network. We assume that the FC is equipped with sufficient energy, hence we are only concerned about the distributed nodes in the set $\mathcal{N}$.

For simplicity, AWGN channel with quasi-static fading is considered. Signal received at the FC from node $j\epsilon\mathcal{N}$ can be expressed as

$$x_j = \hat{\theta}_j + w_j, \tag{6.1}$$

where $w_j$ is the zero-mean receiver noise with variance $\frac{1}{P_j|h_j|^2}$, $P_j$ and $h_j$ are the transmit power of node $j$ and the quasi-static fading coefficient from node $j$ to the FC, respectively. For $N$ nodes, the data vector available at the FC can be written as

$$\mathbf{X}_0 = \begin{pmatrix} \hat{\theta}_0 \\ x_1 \\ \vdots \\ x_N \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \theta + \begin{pmatrix} \tilde{\theta}_0 \\ \tilde{\theta}_1 \\ \vdots \\ \tilde{\theta}_N \end{pmatrix} + \begin{pmatrix} 0 \\ w_1 \\ \vdots \\ w_N \end{pmatrix},$$

so that

$$\mathbf{X}_0 = \mathbf{1}\theta + \tilde{\mathbf{\Theta}}_0 + \mathbf{w}_0. \tag{6.2}$$

where $\tilde{\theta}_i = \hat{\theta}_i - \theta$ is zero-mean with variance $V_i$ for all $i$, $\mathbf{1}$ is the vector of all ones, $\tilde{\mathbf{\Theta}}_0 = [\tilde{\theta}_0, \tilde{\theta}_1, ..., \tilde{\theta}_N]^T$ and $\mathbf{w}_0 = [0, w_1, ..., w_N]^T$. Let us denote by $\Sigma^0 = \Sigma^0_\theta + \Sigma^0_w$ the covariance matrix of $\mathbf{X}_0$, where $\Sigma^0_\theta$ and $\Sigma^0_w$ are the covariance matrices of $\tilde{\mathbf{\Theta}}_0$ and $\mathbf{w}_0$ respectively. The optimal BLUE estimator formed at the FC is given by $\hat{\theta}_{0,n} = \frac{\mathbf{1}^T\left[\Sigma^0\right]^{-1}\mathbf{X}_0}{\mathbf{1}^T[\Sigma^0]^{-1}\mathbf{1}}$ with the updated

estimator variance $V_0^{up} = \left(\mathbf{1}^T \left[\Sigma^0\right]^{-1} \mathbf{1}\right)^{-1}$. The covariance matrix of the observation vector $\mathbf{X}_0$ at the FC is given by (6.3),

$$\Sigma^0 = \begin{pmatrix} V_0 & C_{01} & \dots & C_{0N} \\ C_{10} & V_1 + (P_1|h_1|^2)^{-1} & \dots & C_{1N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{N0} & C_{N1} & \dots & V_N + (P_N|h_N|^2)^{-1} \end{pmatrix}, \tag{6.3}$$

where $C_{ij} = Cov\{\hat{\theta}_i, \hat{\theta}_j\}$ is the covariance between the random variables $\hat{\theta}_i$ and $\hat{\theta}_j$.

## 6.3 Basics of The Shapley Function

For game theoretic formulation of our sequential estimation problem, we consider a game in which the players (sensor nodes) may choose to cooperate by forming coalitions. Cooperative game theory allows competing agents to form coalitions so as to further their individual objectives. In this section, we use the concept of Shapley function [61] as an average measure of fairness for each node.

**Definition 1** A *Shapley function* $\phi(\upsilon)$ is a function that assigns to each possible characteristic function $\upsilon$ a real number, i.e.,

$$\phi\left(\upsilon\right) = \left[\phi_1\left(\upsilon\right), \phi_2\left(\upsilon\right), \cdots, \phi_N\left(\upsilon\right)\right], \tag{6.4}$$

where $\phi_i\left(\upsilon\right)$ represents the worth or value of player $i$ in the game. Note that, the characteristic function of a coalition $S \subset N$ is the *largest* guaranteed payoff to the coalition and is defined as follows:

**Definition 2** Let $2^N$ denote the set of all possible coalitions for the players $N$. Any function $\upsilon : 2^N \rightarrow \mathcal{R}$ satisfying

$$\upsilon(\phi) = 0 \qquad \text{and} \qquad \upsilon(N) \geq \sum_{i=1}^{N} \upsilon(i) \tag{6.5}$$

is a characteristic function of an N-person cooperative game. The Shapley axioms for $\phi(v)$ are [60]:

1. ***Efficiency:*** $\sum_{i\in N}\phi_i(v) = v(N)$.

2. ***Symmetry:*** If $i$ and $j$ are such that $v(S\cup\{i\}) = v(S\cup\{j\})$ for every condition $S$ not containing $i$ and $j$, then $\phi_i(v) = \phi_j(v)$.

3. ***Dummy axiom:*** If $i$ is such that $v(S) = v(S\cup\{i\})$ for every coalition $S$ not containing $i$, then $\phi_i(v) = 0$.

4. ***Additivity:*** If $u$ and $v$ are characteristic functions, then $\phi(u+v) = \phi(v+u) = \phi(u) + \phi(v)$.

It can be proved that there exists a unique function $\phi$ satisfying the above axioms, and this Shapley function can be written as [61]

$$\phi_i = \sum_{S\subseteq\mathcal{N}-i} \frac{(|S|)!\,(N-1-|S|)!}{N!}\left[v(S\cup\{i\}) - v(S)\right]. \tag{6.6}$$

The physical meaning of the Shapely function can be interpreted as follows: suppose that $N$ sensor nodes form a coalition, in which each node joins the coalition in random order. There are $N!$ different ways that the nodes might be ordered in joining the coalition, For any coalition $S$ that does not include node $i$, there are $|S|!\,(N-1-|S|)!$ different ways to order the nodes so that $S$ is the set of nodes who enter the coalition before node $i$. If various orderings are equally likely, $\frac{(|S|)!(N-1-|S|)!}{N!}$ is the probability that, when node $i$ enters the coalition, the coalition $S$ of size $|S|$ is already formed. When node $i$ finds $S$ ahead of it as it joins the coalition, then its marginal contribution to the worth of the coalition is $v(S\cup\{i\}) - v(S)$. Thus under the assumption of randomly-ordered joinings, the Shapley value of each node is its expected marginal contribution when it joins the coalition.

## 6.4 Shapley Value-based Power Allocation

For simplicity of exposition, in this section, we consider the case when all local estimators are uncorrelated and communication is over orthogonal channels. In this chapter, our goal is to estimate a non-random parameter sequentially and at the same time allocate powers among the transmitting/participating nodes in a way that the overall lifetime of the sensor network is maximized. Let $\bar{P}_n$ be the maximum possible transmit power of node $n$ and $P_n$ be the actual transmit power to be obtained from game-theoretic solution. Let us define the lifetime of node $n$ at $\bar{P}_n$ transmit power as $T_n$. Hence according to the definition of lifetime given before, network lifetime, $T_{net} = T_{min} = \min_n T_n$. Then the available power of $n$-th node if all nodes are to have the same $T_{min}$ lifetime is $\psi_n = \frac{\bar{P}_n T_n}{T_{min}}$. We define the coalitional gain of a coalition $S = \{i, j, k\}$ of nodes as:

$$\upsilon(S) = \left\{ \begin{array}{cc} \sum_{n=i,j,k} \lambda_n \psi_n & \text{if} \quad T_i = T_j = T_k \\ \sum_{n=i,j,k} \left(\psi_n - \bar{P}_n\right) & \text{otherwise} \end{array} \right\}, \tag{6.7}$$

where $\lambda_n$ is a suitably chosen weighting parameter which is proportional to the inverse of the variance $V_n$ of the local estimate at node $n$. According to the definition of $\upsilon(S)$ given above, $\upsilon(\emptyset) = 0$. The Shapley value for any node $n$ in the coalition $S$ can be found to be

$$\phi_n = \frac{1}{6} \sum_{\substack{m \neq n \\ m \in S}} \upsilon(\{n, m\}) + \frac{1}{3} \left[\upsilon(S) - \upsilon(S \backslash \{n\})\right]. \tag{6.8}$$

A possible fair allocation of node transmit power can be based on the Shapley value of each node as defined below:

$$\frac{P_j}{P_i} = \frac{\phi_j}{\phi_i} \quad \text{and} \quad \frac{P_k}{P_i} = \frac{\phi_k}{\phi_i}, \tag{6.9}$$

where $P_i$, $P_j$ and $P_k$ are the powers to be committed by the sensor nodes $i$, $j$ and $k$ respectively. The rationale is that a node with higher Shapley value corresponds to having

higher available battery life and better local estimate because of the way the characteristic function $\upsilon(S)$ is formulated. Hence, a node with longer remaining battery life or a better local estimate is allowed to transmit at a higher power than a node with shorter residual battery life, thereby extending the overall lifetime of the sensor network and at the same time achieving the estimation goal.

For the monotonic decrease in the updated variance at the FC, we need $V_0^{up} < V_0$, where $V_0^{up} = \left(\frac{1}{V_0} + \sum_{n=i,j,k} \frac{1}{V_n + (P_n h_n^2)^{-1}}\right)^{-1}$. Let us define $V_0^{up} = \epsilon V_0$ where $0 < \epsilon < 1$. The parameter $\epsilon$ can be used to control the rate at which the updated estimator variance at the FC improves. It also determines the existence of valid solutions for $P_n$'s. The criterion for monotonic decrease of $V_0^{up}$ as more nodes are included in the sequential estimation can be written as

$$\sum_{n=i,j,k} \frac{1}{V_n + (P_n h_n^2)^{-1}} + \frac{1}{V_0}\left(1 - \frac{1}{\epsilon}\right) = 0. \tag{6.10}$$

If $\epsilon$ is too low, the required rate of monotonic decrease in $V_0^{up}$ is too high, and in that case there might not be a feasible solution. On the other hand, if $\epsilon$ is too high, rate of sequential estimation quality at the FC may not be satisfactory. $P_i$, $P_j$ and $P_k$ can easily be solved from (6.9) and (6.10) for the nodes $i$, $j$ and $k$.

We propose the Shapley value based algorithm below to solve for the sequential estimation problem. The details of the algorithm is described in Fig. 6.2. The proposed algorithm is particularly suitable for sensor networks in which it is necessary to control the rate at which FC reaches the target quality of estimation, while at the same time achieving increased network lifetime.

Our proposed algorithm guarantees that when the three nodes satisfy the criteria $V_0^{up} < V_0$, the nodes with longer battery lives transmit at higher power levels. Since FC has the knowledge of the quality of estimates at nodes $i$, $j$ and $k$, all the calculations can be done at the FC. Then it sends control signals to selected nodes informing them the allocated power

**Algorithm 3** Shapley value based Power Allocation

1. FC (node-0) picks any three nodes $i, j, k \in \mathcal{N}$, $i \neq j \neq k$ randomly or in a predetermined order. Calculate $P_i$, $P_j$ and $P_k$ from (6.9) and (6.10).
2. FC sends control signals to nodes $i$, $j$, $k$, and ask them to transmit their signals with the allocated powers. Rest of the nodes are in sleep mode. Update the estimator variance $V_0^{UP}$ at the FC and remove $i$, $j$ and $k$ from the set $\mathcal{N}$.
3. Repeat the same procedures until $V_o^{UP} \leq V_t$ or $\mathcal{N} = \emptyset$ where $V_t$ is the desired quality of estimate at the FC.

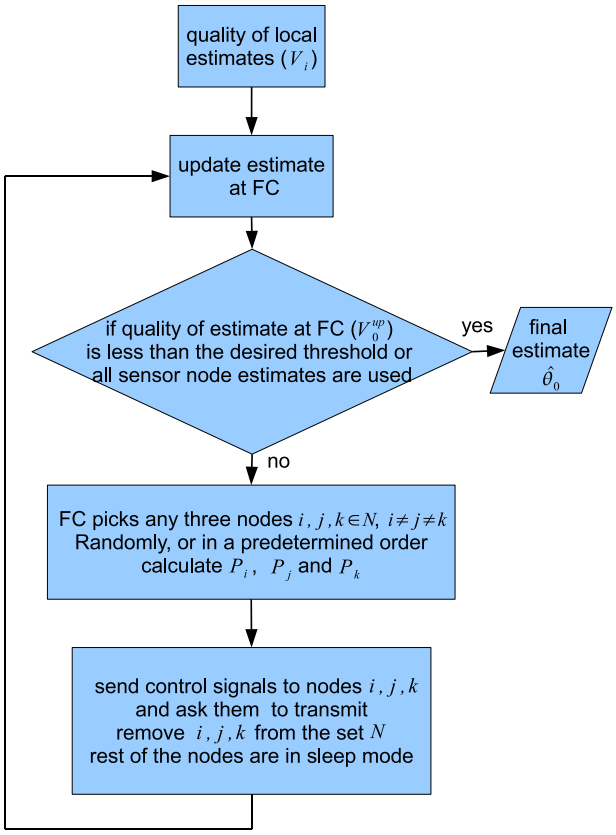



Figure 6.2: Sequential estimation using the proposed algorithm.

levels.

## 6.5 Simulation Results

In this section, we investigate the performance obtained by our proposed algorithm. We will compare the performance with the case when all nodes transmit at their maximum powers. Parameters used for simulations are: number of distributed nodes $N = 51$, estimator variances $V_i \sim \mathcal{U}[1, 20]$. We assume that all channel gains follow Rayleigh distributions with all channel coefficients normalized so that $\mathbb{E}\{h^2\} = 1$. For our simulation, we have chosen $\lambda_n = \frac{V_{max}}{V_n}$, where $V_{max} = \max_n V_n$.

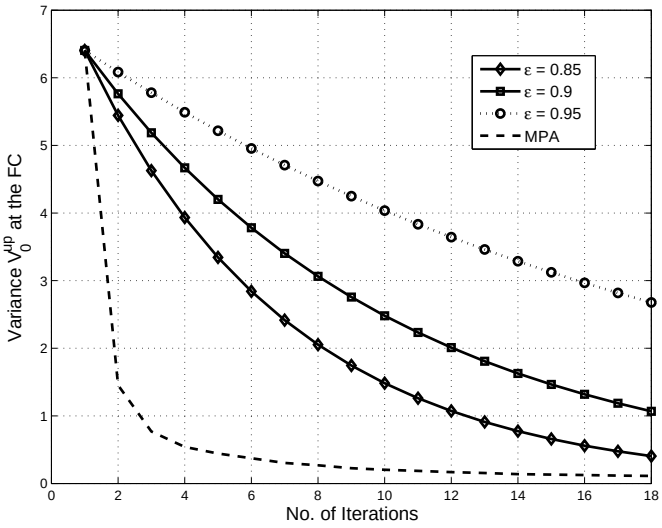



Figure 6.3: Updated variance at the FC using Shapley value-based algorithm for different values of $\epsilon$.

Figure 6.3 shows the performance of the proposed algorithm as local estimates from the distributed nodes are used by the FC to sequentially update its own estimates. For each iteration in Fig. 6.3, local estimates from 3 different nodes are considered. As it can be seen, the quality of estimate at the FC improves as more nodes are incorporated into estimation

process. Algorithm 3 provides the system manager with the flexibility to control the rate at which the variance of estimate at the FC decreases. In Fig. 6.3, we have shown how the values of $\epsilon$ can affect the rate of improvement of the quality of estimate at the FC. The monotonic decrease in variance at the FC center is also shown for the case when all nodes transmit at their maximum powers (MPA) with the proposed algorithm. By MPA (Maximum Power Allocation), we mean to follow all the steps in the proposed algorithm except for setting $P_i = \bar{P}, \forall i \in \mathcal{N}$, where $\mathcal{N}$ is the set of all nodes.

Figure 6.4 shows network lifetime improvement at the end of each time period or estimation block using the proposed algorithms. By each time period or estimation block, we mean that all distributed sensor nodes monitor/sense a PoI and the data collected by the sensor nodes are sent to the FC for estimation purpose. We assume that each distributed node is provided with fixed limited energy at the beginning of the estimation process. After each estimation block, some amount of node energy is dissipated due to data communication and processing, and new lifetime of the sensor network is updated by the FC. At the end of the $j$-th time block, the network lifetime can be defined as $T_{net}^{j} = \min_i \left( \frac{E_i^{j-1} - P_i^j t_p}{\bar{P}} \right)$, where $E_i^{j-1}$ is the energy available at the sensor node $i$ at the end of $(j-1)$-th time period, $P_i^j$ is the power to be spent by the node $i$ during the $j$-th estimation block and $t_p$ is the time each node spends transmitting to the FC. For our simulation, we have used $t_p = 1$ sec, $\bar{P} = 1$ watt and initial node energy $E_i^0 = 10$ joule for $i \in \mathcal{N}$. It can be seen that there is considerable increase in the overall network lifetime using the proposed algorithm. This is because: Algorithm 1 ensures that node with comparatively higher battery life is allowed to transmit at a higher power than that with less battery life, which leads to an overall increase of the network lifetime.

In Fig. 6.5, we assume that sensor nodes' lifetimes are uniformly distributed as $\mathcal{U} \sim [0, 10]$. Network lifetime using our proposed algorithm and MPA are obtained by averaging over 100 initial lifetime realizations. Figure 6.5 shows the average lifetime performance and improvement in the updated variance at the FC as a function of the time period/estimation

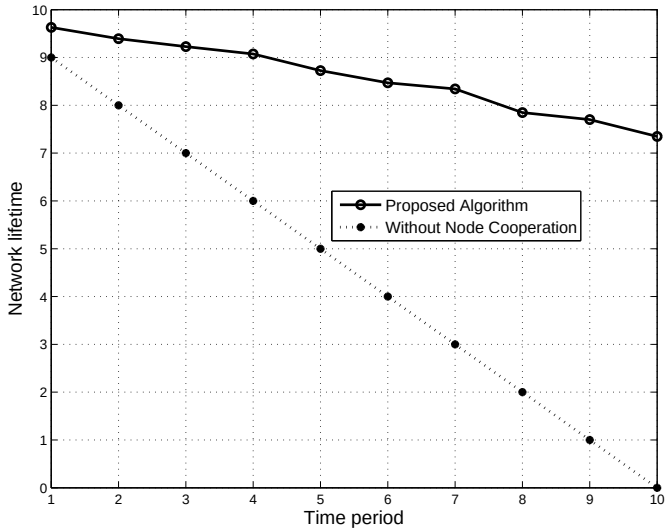
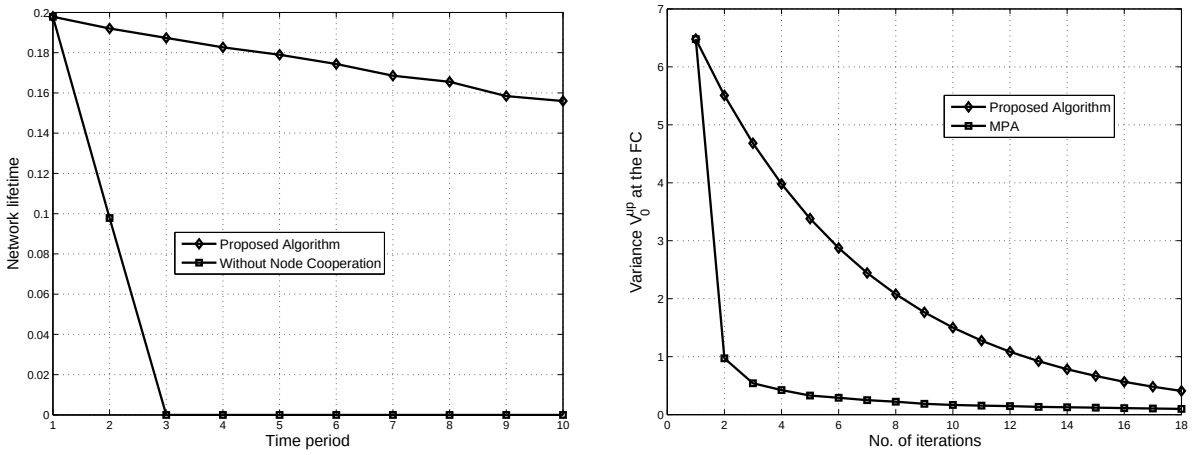



Figure 6.4: Network lifetime improvement using the proposed algorithm for equal initial node energy with $\epsilon = 0.85$.

block. We have used $t_p = 0.1$ sec and $\bar{P} = 1$ watt. It can be seen that although MPA reaches the desired quality of estimate at the FC faster than that with the proposed algorithm, our proposed algorithm outperforms MPA as far as network lifetime improvement is the objective.

Figure 6.6 shows network lifetime as a function of processing time $t_p$. Network lifetime is represented in terms of the number of estimation tasks the sensor network can perform before at least one node runs out of energy. In obtaining Fig. 6.6, we have fixed the rate controlling parameter $\epsilon = 0.85$ and target final variance at the FC $V_t = 0.15\times$(*initial variance at the FC*). At the end of each estimation task node lifetimes are updated by subtracting the energy spent. It can be seen from Fig. 6.6 that the proposed algorithm performs better than the MPA in terms of network lifetime improvement. As one would expect while the processing time $t_p$ increases, network lifetime in both cases decreases. Note that, MPA might achieve the target quality of estimate at the FC with a few nodes transmitting at a higher power while the proposed algorithm might use a large number of nodes transmitting at a lower power levels.




(a) (b)

Figure 6.5: Improvement in the (a) Network lifetime, (b) Variance of estimation at the FC for random initial node energy with $\epsilon = 0.85$.

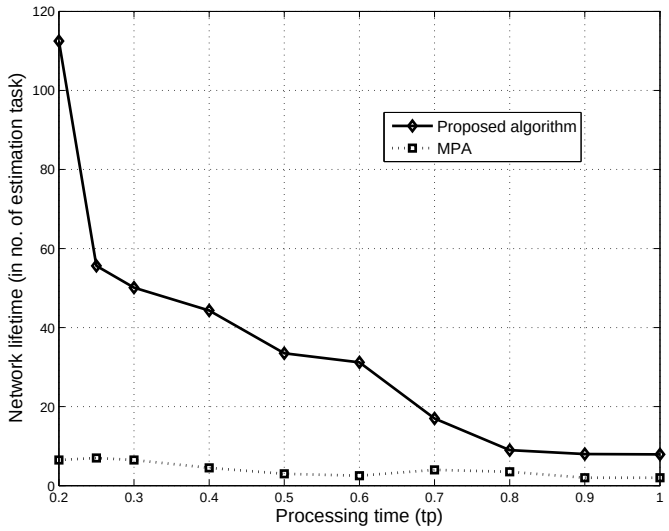



Figure 6.6: Network lifetime in number of estimation task as a function of processing time ($t_p$) for random initial node energy.

Let us also define the lifetime of the sensor network as the time at which a certain fraction $\alpha \in [0, 1]$ of the distributed nodes have the remaining energy/lifetime below a certain threshold value. For our simulation we assume that the initial lifetimes are distributed as

$\mathcal{U} \sim [5, 10]$, processing time $t_p = 2$ sec and lifetime threshold to be 4.8 sec. We have fixed the rate controlling parameter $\epsilon = 0.85$ and target final variance at the FC $V_t = 0.75\times$(*initial variance at the FC*). Node lifetimes are updated at the end of each estimation task. It can be seen from Fig. 6.7 that the proposed algorithm performs better than the MPA in terms of network lifetime improvement. Figure 6.7 shows that for both of the cases, network lifetime increases with the increase of $\alpha$ which is expected. Note that the improvement in network lifetime is more with the proposed algorithm because of the power allocation strategy used in the proposed Algorithm 1.

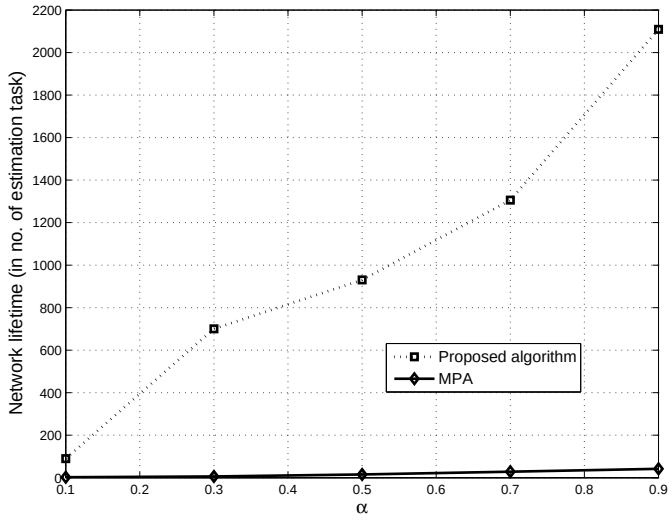



Figure 6.7: Network lifetime in number of estimation task as a function of $\alpha$ for random initial node energy.

## 6.6 Conclusion

In this chapter, a cooperative game-theoretic framework has been proposed to achieve a fair allocation of transmit power for collaborating nodes in a Fusion center (FC) based wireless sensor network tasked with sequential estimation of a non-random parameter. In particular,

we proposed an algorithm based on the concept of the Shapley function to arrive at a fair allocation of power for the nodes. The proposed method offers a system designer with a estimation quality vs. network lifetime trade-off by allowing for tweaking a suitable design parameter. Through simulation results, we have shown that the proposed algorithm achieves target quality of estimate at the FC while improving the overall network lifetime.

# Chapter 7

# Fair Resource Allocation in WSNs for Sequential Estimation

## 7.1 Introduction

Collaborative information processing has a key role to play in efficient handling of large volumes of data collected by mobile sensor nodes in a sensor network. While collaborative information processing for resource-constrained sensor networks has been explored over the last several years [50, 51], still there is a lack of a formal analytical framework for designing collaborative information processing that allows tradeoff between resource and performance. In this chapter, we show that cooperative game theoretic concepts can be applied in approaching such collaborative signal processing problems in WSN. Power management in wireless sensor networks using a non-cooperative game theoretic approach was addressed in [53, 54]. While there is much literature on sequential estimation in WSNs [83, 84], very few attempts have been made to exploit the rich collection of cooperative game theory in power/energy-constrained WSNs tasked with estimating a Phenomenon of Interest (PoI).

In this chapter, the problem of sequential estimation in a WSN is formulated in a coop-

erative game theoretic framework, in particular, using Nash Bargaining [57]. While most of the existing literature mentioned in section [50, 51, 53, 54, 83, 84] deal the power allocation and parameter estimation as separate problems, in this chapter we formulate and solve the problems of sequential estimation and power allocation in WSNs as a combined problem. We propose a simple Nash bargaining based solution to the problem of power allocation for sensing nodes that are subjected to power constraints.

The remainder of this chapter is organized as follows: Section 7.2 presents the assumed sensor network model for the sequential estimation problem. Section 7.3 introduces the bargaining problem and summarizes the Nash bargaining solution (NBS). In Sections 7.4 and 7.5, we introduce two fair power allocation algorithms based on the NBS. Section 7.6 evaluates the performance of the proposed algorithms via simulations. Finally, Section 7.7 concludes the chapter by summarizing our results.

## 7.2 Sensor Network Model for Sequential Estimation

In this chapter, we use the same system model described in section 6.2 except for different representation of the received signal. Similar as before, we assume that the FC updates its estimator sequentially by combining its own observation with the noise-corrupted estimators of the other selected nodes received over noisy communication links. We consider the case when the FC does not want the distributed nodes to transmit at their maximum powers, while achieving a target estimator quality using a few number of nodes, and/or less total consumed power possible. For simplicity, AWGN channel with quasi-static fading is considered. Signal received at the FC from node $j \epsilon \mathcal{N}$ can be expressed as

$$x_j = \sqrt{\alpha_j}\hat{\theta}_j + w_j, \tag{7.1}$$

where $w_j$ is the zero-mean receiver noise with variance $\frac{\nu_j^2}{P|h_j|^2}$, $\alpha_j = \frac{P_j}{P}$ is the ratio of the transmit power $P_j$ of node $j$ to the maximum allowed transmit power $P$, and $h_j$ is the fading

coefficient between the communication link from node $j$ to the FC. For $N$ nodes, the data vector available at the FC can be written as

$$\mathbf{X}_0 = \begin{pmatrix} 1 \\ \sqrt{\alpha_1} \\ \vdots \\ \sqrt{\alpha_N} \end{pmatrix} \theta + \begin{pmatrix} \tilde{\theta}_0 \\ \sqrt{\alpha_1}\tilde{\theta}_1 \\ \vdots \\ \sqrt{\alpha_N}\tilde{\theta}_N \end{pmatrix} + \begin{pmatrix} 0 \\ w_1 \\ \vdots \\ w_N \end{pmatrix}, \tag{7.2}$$

so that $\mathbf{X}_0 = \mathbf{a}_0\theta + \tilde{\mathbf{\Theta}}_0 + \mathbf{w}_0$, where $\tilde{\theta}_i = \hat{\theta}_i - \theta$ is zero-mean with variance $V_i$ for all $i$, $\mathbf{a}_0 = [1, \sqrt{\alpha_1}, \cdots, \sqrt{\alpha_N}]^T$, $\tilde{\mathbf{\Theta}}_0 = [\tilde{\theta}_0, \sqrt{\alpha_1}\tilde{\theta}_1, \cdots, \sqrt{\alpha_N}\tilde{\theta}_N]^T$ and $\mathbf{w}_0 = [0, w_1, \cdots, w_N]^T$. Let us denote by $\Sigma^0 = \Sigma^0_\theta + \Sigma^0_w$ the covariance matrix of $\mathbf{X}_0$, where $\Sigma^0_\theta$ and $\Sigma^0_w$ are the covariance matrices of $\tilde{\mathbf{\Theta}}_0$ and $\mathbf{w}_0$ respectively. Thus, the optimal BLUE estimator formed at the FC is given by $\hat{\theta}^1_0 = \frac{\mathbf{a}_0^T\left[\Sigma^0\right]^{-1}\mathbf{X}_0}{\mathbf{a}_0^T[\Sigma^0]^{-1}\mathbf{a}_0}$ with the updated estimator variance $V_0^1 = \left(\mathbf{a}_0^T\left[\Sigma^0\right]^{-1}\mathbf{a}_0\right)^{-1}$.

## 7.3 Basics of Nash Bargaining Solution

In this section, we briefly introduce the concept of Nash bargaining solution (NBS) [60] and then apply it to achieve a fair allocation of power among nodes in our sensor network.

The bargaining problem in a cooperative game can be described as follows [60]: Let $\mathcal{N} = \{1, 2, \cdots, N\}$ be the set of players, and let $\mathbf{S}$ be a closed and convex subset of $\mathcal{R}^N$ representing the set of feasible payoff allocations that the players can get if they all cooperate. Let $u^i_{min}$ be the minimum expected payoff for the $i$-th player, below which it will not cooperate. Suppose $\{u_i \in \mathbf{S} | u_i \geq u^i_{min}, \forall i \in \mathcal{N}\}$ is a nonempty bounded set. Define $\mathbf{u}_{min} = (u^1_{min}, \ldots, u^N_{min})$; then the pair $(\mathbf{S}, \mathbf{u}_{min})$ is called the $N$-person bargaining problem. NBS provides a unique and fair Pareto optimal point under the conditions given in [60].

**Theorem 4** (*Existence and Uniqueness of NBS*) [60]: If $\bar{\mathbf{u}}$ is said to be an NBS in $\mathbf{S}$ for $\mathbf{u}_{min}$, i.e., $\bar{u} = \psi(\mathbf{S}, \mathbf{u}_{min})$, then the solution function $\psi(\mathbf{S}, \mathbf{u}_{min})$ is unique, and the following axioms are satisfied:

1. ***Individual rationality:*** $\bar{u}_i \geq \mathbf{u}_{min}^i, \forall i$.
2. ***Feasibility:*** $\bar{\mathbf{u}} \in \mathbf{S}$.
3. ***Pareto optimality:*** For every $\hat{\mathbf{u}} \in \mathbf{S}$, if $\hat{u}_i \geq \bar{u}_i, \forall i$, then $\hat{u}_i = \bar{u}_i, \forall i$.
4. ***Independence of irrelevant alternatives:*** If $\bar{\mathbf{u}} \in \mathbf{S}' \subset \mathbf{S}$, $\bar{\mathbf{u}} = \psi(\mathbf{S}, \mathbf{u}_{min})$, then $\bar{\mathbf{u}} = \psi(\mathbf{S}', \mathbf{u}_{min})$.
5. ***Independence of linear transformations:*** For any linear scale transformation $\Psi$, $\Psi[\psi(\mathbf{S}, \mathbf{u}_{min})] = \psi[\Psi(\mathbf{S}), \Psi(\mathbf{u}_{min})]$.
6. ***Symmetry:*** If $\mathbf{S}$ is invariant under all exchanges of agents, $\psi_j(\mathbf{S}, \mathbf{u}_{min}), \forall j, j'$.

And it also satisfies

$$\psi(\mathbf{S}, \mathbf{u}_{min}) \in \arg \max_{\bar{\mathbf{u}} \in \mathbf{S}, \bar{u}_i \geq u_{min}^i, \forall i} \prod_{i=1}^{N} \left(\bar{u}_i - u_{min}^i\right)^{\beta_i}, \tag{7.3}$$

where $\beta_i$ is the bargaining weight associated with the payoff of player $i$. Intuitively, it means how much importance is given to a particular player in the bargaining process.

The cooperative game for the sequential estimation problem can be described as follows: Each player (sensor) has $u_i$ as its objective function, where $u_i$ is non-negative, bounded from above and has a nonempty, closed and convex support. The goal is to maximize all $u_i$'s simultaneously. $u_{min}^i$ is the minimal payoff that player $i$ would obtain if it had not cooperated with other players.

## 7.4 NBS-based Solution Using Algorithm 4

For simplicity of exposition, in this section, we consider the case when all local estimators are uncorrelated and communication is over orthogonal channels. Let us assume that $\nu_i^2 = \sigma^2$

for $i \in \mathcal{N}$. Hence the covariance matrix at the FC can be written for node $i$ and $j$, as $\Sigma^0 = diag\left(V_0, \alpha_i V_i + \frac{\sigma^2}{P|h_i|^2}, \alpha_j V_j + \frac{\sigma^2}{P|h_j|^2}\right)$.

In this chapter, the utility of a node $i$ is defined as the inverse of the quality of its local estimator: $u_i = \frac{1}{V_i}$. Hence, the optimization goal is to determine nodes $i$ and $j$'s transmission powers to the FC, with the precondition that none of these nodes transmit at their full powers. To that end, we proposed the following objective function to be maximized:

$$U \;=\; u_0 - u_{min}^0 = \frac{\alpha_i}{\alpha_i V_i + \frac{\sigma^2}{P|h_i|^2}} + \frac{\alpha_j}{\alpha_j V_j + \frac{\sigma^2}{P|h_j|^2}}, \tag{7.4}$$

where $u_0$ is the utility corresponding to the updated estimator at the FC when node $i$ and $j$ share their estimators with node 0, and $u_{min}^0$ is the minimum possible payoff of node-0 that it would expect from the bargaining process. Thus, the optimization problem is

$$\min_{\alpha_i, \alpha_j} -U \qquad \text{s.t.} \qquad \left\{ \begin{array}{r} -\alpha_i \leq 0, -\alpha_j \leq 0 \\ \alpha_i + \alpha_j - 1 \leq 0 \end{array} \right\}. \tag{7.5}$$

Note that, we have defined the power constraint to be $\alpha_i + \alpha_j \leq 1$. The assumption is reasonable because we do not want node $i$ and $j$ to transmit at their maximum powers. Thus, we would like to allocate the total power $P$ between nodes $i$ and $j$ in a fair way (in this case, using NBS) for sequentially estimating the parameter $\theta$ provided that no node transmits at its maximum power. To that end, we propose the following NBS based Algorithm 4 to solve the power allocation problem. The proposed sequential estimation process is summarized in Algorithm 4 and described in detail in Fig. 7.1.

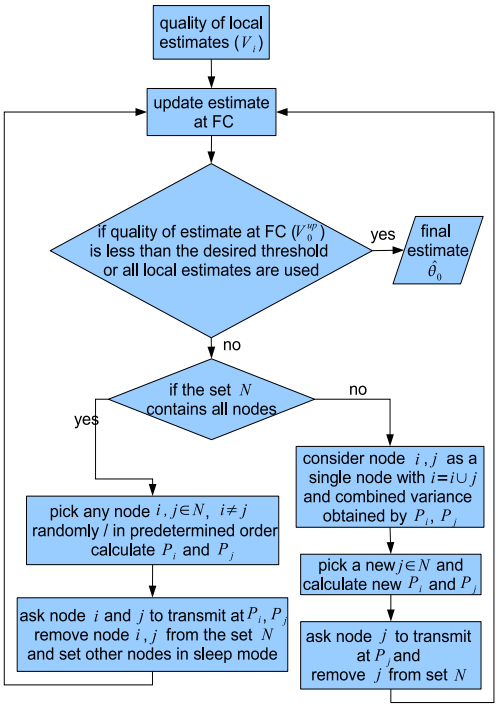



Figure 7.1: Sequential estimation process using Algorithm 4.

**Algorithm 4** NBS based Power Allocation

1. FC picks any two nodes $i, j \in \mathcal{N}$, $i \neq j$ randomly or in a predetermined order.
2. Calculate $\alpha_i$ and $\alpha_j$ from (7.6). Assign powers $P_i$ and $P_j$ to nodes $i$ and $j$ respectively. Update the estimator variance at FC $V_0^{up}$ and remove $i$ and $j$ from the set $\mathcal{N}$.
3. Consider nodes $i$ and $j$ as a single node with combined variance obtained by using $\alpha_i$ and $\alpha_j$, and assign $i = i \cup j$. FC picks a new node $j$ from the set $\mathcal{N}$ and calculate new $\alpha_i$ and $\alpha_j$. Assign power $P_j$ to new node $j$ and update the estimator variance at FC $V_0^{up}$.
4. Repeat the same procedures from step 3 until $V_o^{UP} \leq \epsilon$ or $\mathcal{N} = \emptyset$, where $\epsilon$ is the desired quality of estimate at the FC.

Since $U$ is concave in $\alpha_i$, $\alpha_j$ and optimization constraints are linear, the Karush-Kuhn-Tucker (KKT) [85] conditions are both necessary and sufficient. Solving for $\alpha_i$ and $\alpha_j$ from

(7.5), we get the solution:

$$\left.\begin{aligned}\alpha_i &= \frac{V_j|h_j| + \frac{\sigma^2}{P}\left(\frac{1}{|h_j|} - \frac{1}{|h_i|}\right)}{V_i|h_i| + V_j|h_j|},\\ \alpha_j &= \frac{V_i|h_i| + \frac{\sigma^2}{P}\left(\frac{1}{|h_i|} - \frac{1}{|h_j|}\right)}{V_i|h_i| + V_j|h_j|}.\end{aligned}\right\} \tag{7.6}$$

For $|h_i|^2 = |h_j|^2$, $\alpha_i = \frac{V_j}{V_i+V_j}$ and $\alpha_j = \frac{V_i}{V_i+V_j}$. Hence, node with more accurate estimation are allowed to transmit at a higher power than that with less accurate estimation, which intuitively makes sense. Since FC has the knowledge of the quality of estimates at nodes $i$ and $j$, all the calculations can be done at the FC and it can send control signals to nodes $i$ and $j$ to transmit at powers $P_i = \alpha_i P$ and $P_j = \alpha_j P$ respectively.

## 7.5 NBS-based Solution Using Algorithm 5

It is to be noted that for the NBS-based Algorithm 4 above, an explicit analytical solution in the case of correlated observations could not be obtained. This motivated us to propose an NBS-based Algorithm 5 to solve the problem of power allocation for collaborating nodes with correlated observations. We assume that all local estimators are correlated such that $C_{ij} = Cov\{\hat{\theta}_i, \hat{\theta}_j\} = \rho, \forall i, j$, where $C_{ij}$ is the covariance between the random variables $\hat{\theta}_i$ and $\hat{\theta}_j$. The covariance matrix at the FC and at node $i$ (if $i$-th node were to receive signal from the FC) can be written respectively as: $\Sigma^0 = \begin{pmatrix} V_0 & \sqrt{\alpha_i}\rho \\ \sqrt{\alpha_i}\rho & \alpha_i V_i + \frac{\sigma^2}{P|h_i|^2} \end{pmatrix}$, and $\Sigma^i = \begin{pmatrix} V_i & \sqrt{\alpha_0}\rho \\ \sqrt{\alpha_0}\rho & \alpha_0 V_0 + \frac{\sigma^2}{P|h_i|^2} \end{pmatrix}$ for $i \in \mathcal{N}$. Hence we have

$$u_0 - u^0_{min} = (1\ \sqrt{\alpha_i})[\Sigma^0]^{-1}(1\ \sqrt{\alpha_i})^T - \frac{1}{V_0}, \tag{7.7}$$

$$\text{and} \quad u_i - u^i_{min} = (1\ \sqrt{\alpha_0})[\Sigma^i]^{-1}(1\ \sqrt{\alpha_0})^T - \frac{1}{V_i}. \tag{7.8}$$

As a result, the optimization problem becomes:

$$\min_{\alpha_0,\alpha_i} -\prod_{j=0,i}\left(u_j - u_{min}^j\right) \text{ s.t. } \left\{ \begin{array}{c} -\alpha_0 \leq 0, -\alpha_i \leq 0 \\ \alpha_0 + \alpha_i - 1 \leq 0 \end{array} \right\} \tag{7.9}$$

Since the optimization problem is again convex in $\alpha_0$, $\alpha_i$ and constraints are linear, the KKT conditions are again both necessary and sufficient. Solving for $\alpha_i$ from (7.9), the only non-zero solutions that satisfy all the KKT conditions can be obtained as:

$$\alpha_i = \left[\frac{1 + \mathcal{Q} \mp \sqrt{1+\mathcal{Q}}}{\mathcal{Q}}\right]^+, \tag{7.10}$$

where $\mathcal{Q} = \frac{(V_0 - V_i)(V_0 V_i - \rho^2)}{V_i\left(V_0 \frac{\sigma^2}{P|h_i|^2} + V_0 V_i - \rho^2\right)}$ with $V_0 > 0$, $V_i > 0$ and $0 \leq \rho \leq 1$, and $[.]^+$ means only the non-negative bounded values are considered. It is to be noted that the FC only helps to calculate $\alpha_i$ and it discards the value $\alpha_0$, as it only transmits control signals. Since all the calculations are done at the FC, the only parameters the FC needs to know are $V_i$ and $\frac{\sigma^2}{P|h_i|^2}$. The details of the sequential estimation process using NBS-based Algorithm 5 is described in Fig. 7.2.

**Algorithm 5** NBS based Power Allocation

1. FC (node-0) picks any node $i \in \mathcal{N}$, randomly or in a predetermined order.
2. Calculate $\alpha_i$ from (7.10). Assign power $P_i$ to node $i$. Update the estimator variance at FC $V_0^{up}$ and remove node $i$ from the set $\mathcal{N}$.
3. Repeat the same procedures until $V_o^{up} \leq \epsilon$ or $\mathcal{N} = \emptyset$.

## 7.6 Simulation Results

In this section, we investigate the performance obtained by our proposed algorithms. Parameters used for simulations are: number of distributed nodes $N = 50$, estimator variances

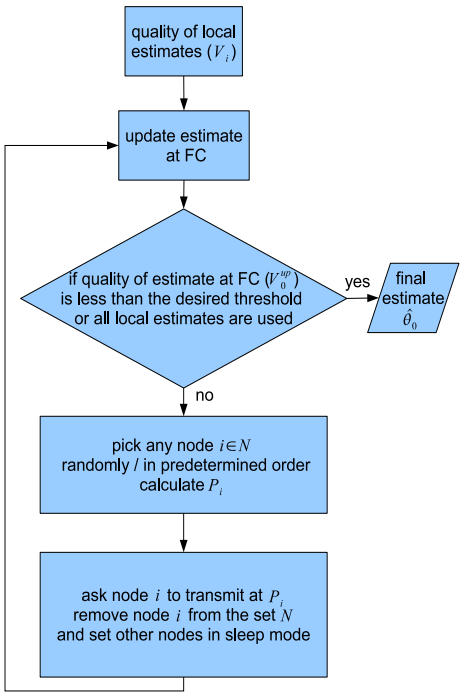



Figure 7.2: Sequential estimation process using Algorithm 5.

$V_i \sim \mathcal{U}[1, 20]$, and $\frac{\sigma^2}{P} = 0.5$. We assume that all channel gains follow Rayleigh distributions with all channel coefficients normalized so that $\mathbb{E}\{h^2\} = 1$. All simulation results are obtained by averaging over 500 fading realizations.

Figure 7.3 shows the updated estimator variance $V_0^{up}$ at the FC as a function of number of nodes. As it can be seen from Fig. 7.3, $V_0^{up}$ monotonically decreases for both algorithms as more nodes are incorporated into the estimation process. The monotonic decrease in variance at the FC center is almost as good as the case when all the nodes transmit at their maximum powers (MPA) with each of the proposed algorithms. By MPA (Maximum Power Allocation), we mean to follow all the the steps in each of the proposed algorithms except for setting $\alpha_i = 1, \forall i$. As it can be seen from Fig. 7.3, Algorithm 4 outperforms Algorithm 5 in terms of the rate of improvement of the estimator quality at the FC. This is because

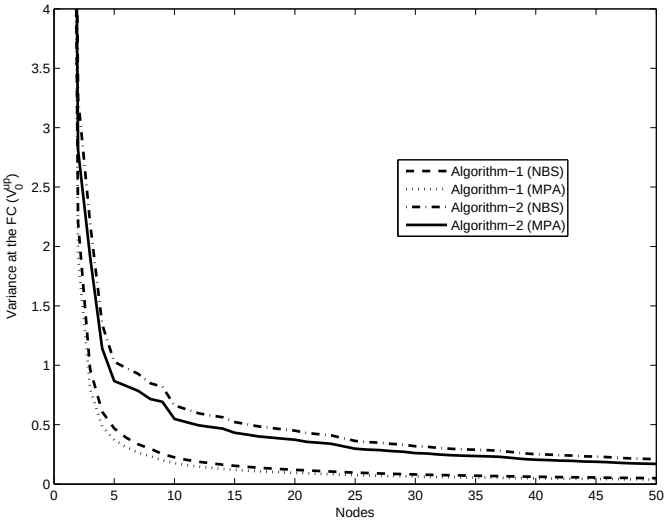



Figure 7.3: Updated variance at the FC vs. number of nodes with $\rho = 0$.

of the formation of Nash product (NP) of the optimization problems in (7.5) and (7.9). NP in (7.5) was obtained by setting $\beta_0 = 1$ and $\beta_i = \beta_j = 0$. Hence, the emphasis is only on maximizing the payoff of node 0 (FC). On the other hand, in (7.9), we have set $\beta_0 = 1$ and $\beta_i = 1$. As a result, the optimization problem tries to maximize the payoffs of both nodes 0 and $i$. Since the payoff of a node is defined as the inverse of the quality of estimate, better rate of improvement of the estimator quality is achieved by maximizing the payoff.

Figure 7.4 shows network lifetime improvement at the end of each time period or estimation block using the proposed algorithms. Note that, we define the lifetime of a sensor network as the time after which at least one or a certain fraction of sensor nodes run out of their batteries, resulting in a hole within the network. Since FC is equipped with sufficient energy, we are only concerned about the distributed nodes. We assume that each distributed node is provided with limited energy at the beginning of the estimation process. At the end of the $j$-th time block, the network lifetime can be defined as $T_N^j = \min_i \left( \frac{E_i^{j-1} - P_i^j t_p}{P} \right)$, where $E_i^{j-1}$ is the energy available at the sensor node $i$ at the end of $(j-1)$-th time period, $P_i^j$ is the power to be spent by the node $i$ during the $j$-th estimation block and $t_p$ is the time each

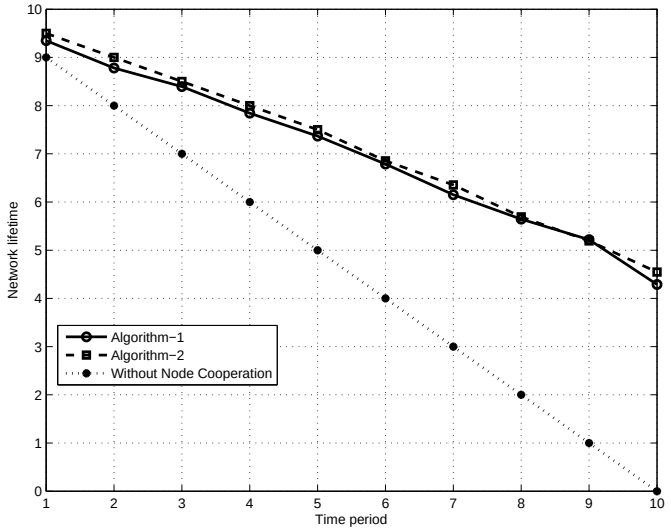



Figure 7.4: Network lifetime improvement using Algorithm 4 and 5.

node spends transmitting to the FC. For our simulation, we have used $E_i^0 = 10$ Joules, $t_p = 1$ sec and $P = 1$ watt. As it can be seen from Fig. 7.4, Algorithm 5 is slightly better than the Algorithm 4 as far as network lifetime improvement is concerned. This is again because of the formation of the optimization problems for Algorithms 4 and 5: the obtained value of $\max_i \alpha_i$ for Algorithm 4 can be higher than that for the Algorithm 5 most of the time, which reduces the lifetime of WSN using Algorithm 4 compared to that with Algorithm 5.

## 7.7 Conclusion

In this chapter, a NBS based framework has been proposed to achieve a fair allocation of transmit power for collaborating nodes in a FC-based wireless sensor network tasked with sequential estimation of a non-random parameter. In particular, we proposed two algorithms based on the concept of Nash bargaining to arrive at a fair allocation of power for the nodes. Simulation results show that the proposed algorithms sequentially achieve the desired quality of estimate at the FC, and increase the overall network lifetime.

# Chapter 8

# Summary of Dissertation and Future Works

In this dissertation, we proposed parametric and nonparametric-based diffusive source localization and tracking schemes using wireless sensor network. We also addressed research challenges in a resource constrained sensor network tasked with estimating a desired PoI. In this chapter, we summarize our contributions by highlighting the main results, and further discuss some of the possible research directions that this work can be extended to.

## 8.1 Summary of Results

In Chapter 2, we derived the space-time concentration distribution for a special diffusion phenomenon, i.e., an underwater oil spill, by modeling and solving corresponding diffusion equations with appropriate initial, boundary conditions, and pragmatic assumptions. The physical modeling and solution techniques used in this chapter, are applicable (with appropriate modifications) to many other similar contexts of practical importance as well. Our main objective was to obtain an analytical solution to the diffusion problem, rather than

using non-model based sophisticated numerical techniques.

In Chapter 3, two parametric estimation methods based on the MLE and BLUE for estimating static diffusive source location are proposed. We obtained the Cramér-Rao lower bound as theoretical performance bound for source localization. The pros and cons of the two proposed methods are also mentioned. While the MLE requires complete knowledge of the pdf, knowing only 1st and 2nd order information of data are enough for BLUE-based technique. On the other hand, since BLUE can only be used for linear observation models, hence linearization is necessary before using BLUE for parameter estimation. It was observed that performance of the MLE is slightly better than the CRLB in some cases. This is an example of a typical bias-variance trade-off, and therefore MLE outperform the CRLB by trading variance for bias. As one would expect, the MLE performs better than the BLUE-based diffusive source localization method. However, the later shows satisfactory performance trend for large number of sensing nodes and time samples.

Although performance of our proposed parametric source localization methods in Chapter 3 were demonstrated to be satisfactory, they do, however, come with an important drawback, i.e., knowledge of complete parametric description of the underlying statistical model. In many complex practical problems, complete knowledge of the likelihood function, is sometimes, a strict requirement. To that end, in Chapter 4, we proposed novel non-parametric diffusive source localization methods based on DPMM for both single and multiple diffusive sources. We exploited the excellent features of DPMM as a classifier, for our source estimation problem. We analytically proved convergence of the proposed algorithms in terms of total variation norm. For the proposed algorithms, knowledge of the family of distribution of the likelihood function is enough, contrary to the parametric estimation technique that requires complete and suitable description of the system model and the likelihood function. It was also observed that the proposed algorithms offer low computational complexity for implementation.

In Chapter 5, we developed a particle filter based target tracking method for moving

diffusive source in a dispersive medium. The PCRLB corresponding to moving diffusive source tracking was obtained as a theoretical performance measure and compared with the simulation results. The quality of the tracking performance with the proposed tracking scheme was found to be satisfactory. A recursive formula for the derivation of the PCRLB is presented. Both grid-based and random node deployment strategies were investigated for our proposed tracking scheme. The effect of sampling time on the moving source tracking was also studied.

In Chapter 6, a sequential methodology for parameter estimation along with fair power allocation is developed from game-theoretic perspective. The Shapley value based cooperative game-theoretic framework has been proposed to achieve a fair allocation of transmit power for collaborating nodes in a Fusion center (FC) based wireless sensor network tasked with sequential estimation of a non-random parameter. The proposed method offers a system designer with a estimation quality vs. network lifetime trade-off by allowing for tweaking a suitable design parameter. We have shown that the proposed algorithm achieves target quality of estimate at the FC while improving the overall network lifetime through numerical simulations.

Finally, in Chapter 7, NBS based framework has been proposed to achieve a fair allocation of transmit power for collaborating nodes in a FC-based wireless sensor network tasked with sequential estimation of a non-random parameter. In particular, we proposed two algorithms based on the concept of Nash bargaining to arrive at a fair allocation of power for the nodes. The novelty of the proposed framework is that it combines the estimation and resource allocation problem together, and provides a single solution to the sequential estimation and fair power allocation problems.

## 8.2 Future Work

This dissertation can lead to the following possible future research directions:

*Modeling of more complex and realistic diffusion phenomena*: In Chapter 2, we discussed the physical modeling and mathematical formulation of a special diffusion phenomenon, i.e., an underwater oil spill scenario, and obtained the spatio-temporal concentration distribution. In our formulation, we considered laminar water velocity as the turbulent effect [86]. For realistic scenario, one should also consider the external velocity acting in the diffusion field including the turbulence effects which is highly chaotic in nature, also space and time dependant. Hence, the concept of stochastic diffusion and high performance numerical techniques will come handy for this type of transport and turbulence modeling. In future research, it is interesting to combine our obtained analytical results with non-model based numerical techniques to make them applicable for more realistic and complex scenarios.

*Detection or estimation in the presence of intelligent threats*: In Chapter 3 and 4, we investigated parametric and non-parametric diffusive source localization approaches. The techniques can be further modified and extended to detect or estimate smart/intelligent threats. One problem that could be challenging for a WSN is when an agent behind the threat tries to deceive the WSN by strategically deploying different types of diffusive sources, i.e., impulsive vs continuous sources, or sources of different chemical compounds. In these scenarios, if the WSN tries to locate or track either of the sources based on concentration measurements, then the WSN could be deceived by the perturbations caused by the other sources. Another possible threatening strategy could be the generation of *virtual diffusive sources* by combining the effects of several different sources. For example, we can recall the Sarin gas attack that happened on the Tokyo subway in 1995 [87], in which the perpetrator used Sarin ($[(CH_3)_2CHO]CH_3P(O)F$) liquid which can be made by chemical reaction between Methylphosphonyl Difluoride and a mixture of isopropyl alcohol:

$$CH_3P(O)F_2 + (CH_3)_2CHOH \longrightarrow [(CH_3)_2CHO]CH_3P(O)F + HF$$

Hence, there is always some possibility that some chemical compounds which may not be harmful alone, but when they are mixed in right proportion with each other, then they could react to produce chemical threats. These kind of problems can possibly be tackled by

making the WSN smart enough to be able to classify and identify the type of the sources in the surrounding region. This could be achieved by using sophisticated machine learning algorithms at the WSN so that the network can learn and adapt in such situations.

*Effect of realistic channel conditions on WSNs for diffusive source estimation*: Even though a static sensor network meets the performance criteria at the initial deployment stage, it may not adapt to the highly unpredictable dynamics in network conditions such as coverage holes caused by node failures and changing dynamics of the phenomenon being sensed over space and time. For instance, in case of an underwater WSN, it is interesting to study the performance of the proposed source localization and tracking methods in the presence of the problems and requirements for underwater wireless communications. Therefore, it is an important and interesting problem to consider the realistic channel scenarios based on the diffusive field in question so that the effect of the medium-specific conditions are reflected on the obtained solution.

Similar studies can be done in case of our proposed solutions to sequential parameter estimation and fair power allocation for resource constrained WSNs in Chapter 6 and 7. In such scenarios, it would be interesting to see the performance trend when the effect of imperfect channel conditions, such as, channel connectivity, coverage, signal attenuation etc. are taken into account.

# Chapter 9

# Other Research Work Done in Cognitive Radios

## 9.1 Efficient Dynamic Spectrum Sharing in Cognitive Radio Networks: Centralized Dynamic Spectrum Leasing (C-DSL)

### 9.1.1 Introduction

As wireless applications are becoming more widely used, demand for bandwidth is also expected to increase in future years. Under the long-adhered regulatory framework, spectrum appears to be a scarce resource. On the other hand, it has been observed that the perceived scarcity of radio spectrum is mainly due to the inefficiency of traditional spectrum allocation policies [88, 89]. This led the FCC to recommend three broad solutions to improve the spectrum utilization in its 2002 Spectrum Policy Task Force Report: a) spectrum reallocation, b) spectrum leasing, and c) spectrum sharing. The first of these was meant to be a

long-term solution. Perhaps, the best example is the opening of the 700MHz TV band for cognitive radio operation. Spectrum leasing in [88] was mostly interpreted to be a static, or off-line, solution, at least according to the current literature. As an alternative to the traditional static spectrum management policy, the *dynamic spectrum sharing* (DSS) in [90–94] is considered as an effective way to improve inefficient static spectrum utilization by allowing secondary users to dynamically access the so-called white spaces in spectrum already licensed to the primary users. Some of these spectrum sharing proposals can be identified as being hierarchical-access methods, in that there is usually a primary system that owns the spectrum rights and a secondary system that is interested in accessing this spectrum whenever possible [52, 95]. In almost all existing hierarchical spectrum sharing proposals, the burden of interference management and coexistence is squarely placed on the secondary system. As in [96], we term these proposals as *dynamic spectrum access* (DSA). Cognitive radios, which can be defined as smart radios with built in cognition [97], are especially suited for realizing such dynamic spectrum sharing due to their ability to assess, learn from and orient to the observed RF environment.

Recently [96,98–101] introduced the concept of *dynamic spectrum leasing* (DSL) as a new paradigm for spectrum sharing in cognitive radio networks. The authors identified that the passive primary systems/users that are oblivious to the existence of secondary users is incomplete at the best, and inefficient at the worst, if the objective is to achieve efficient spectrum utilization via DSS. In [98, 99], the primary users were allowed to dynamically manage the interference they experience from the secondary transmissions by adapting their interference cap (IC) according to the observed RF environment and required Quality-of-Service (QoS). Simultaneously, the secondary users aim to achieve energy efficient transmissions, while not causing excessive interference to the primary users. In this chapter, we extend the DSL framework for spectrum sharing by introducing the *centralized dynamic spectrum leasing* (C-DSL) as a new game theoretic model for dynamic spectrum sharing in cognitive radio networks. In particular, we allow for multiple primary users to be simultaneously present in the primary frequency band of interest. In the proposed C-DSL based DSS networks all

primary users participate in the spectrum sharing process as a single system under a central control. We develop an alternative game-theoretic framework for C-DSL based spectrum sharing by identifying new payoff functions for both primary system and secondary users that are motivated by network spectrum utilization considerations. We introduce a general structure for a suitable class of utility functions for the primary system that reflects the demand for spectrum access from the secondary users, the primary system QoS requirements, and analyze the conditions for reaching a desired equilibrium under greedy adaptations. We also generalize the proposed non-cooperative C-DSL game to allow for linear multiuser detectors, in particular the matched filter (MF) and the linear minimum mean squared error (LMMSE) receivers, at the secondary base stations, and establish the existence of an equilibrium in this primary-secondary spectrum-leasing game. Several DSS radio networks based on the proposed C-DSL framework are investigated via simulations to analyze the equilibrium behavior and to identify design guidelines. As in previous work on DSL, we emphasize the need to minimize the need for conscious effort by the primary system to exchange inter-system control information. Indeed, as we will show later, the proposed C-DSL can be implemented with the same two broadcast parameters from the primary system assumed in [96, 99]. The robustness of the C-DSL based spectrum sharing to time-varying fading is also investigated.

Contributions of this chapter that distinguishes it from previous literature are as follows: (i) a novel game-theoretic approach *centralized dynamic spectrum leasing* (C-DSL) is proposed for dynamic spectrum sharing in the presence of multiple primary users extending the model in [99], (ii) we generalize the primary system utility function defined in [96, 99] and introduce a new utility function for the secondary users that leads to efficient utilization of the network spectrum, (iii) the proposed non-cooperative C-DSL game is generalized to allow for linear multiuser detectors, such as the linear minimum mean squared error (LMMSE) receivers at the secondary receivers, and (iv) the robustness of the proposed C-DSL framework is investigated in the presence of slow time varying fading.

The remainder of this chapter is organized as follows: Section 9.1.2 describes the C-DSL

based spectrum leasing cognitive radio network model. Section 9.1.3 presents the proposed game-theoretic model for C-DSL in a DSS based cognitive radio network. Sections 9.1.4 and 9.1.5 discuss the existence of unique Nash equilibria under MF and LMMSE receivers, respectively. Section 9.1.6 evaluates the performance of a spectrum sharing network based on the proposed C-DSL under various conditions and discusses the performance trends and design guidelines. Finally, Section 9.1.7 concludes the chapter by summarizing our results.

### 9.1.2 C-DSL-based Spectrum Sharing Cognitive Radio Network Model

We assume there is one primary wireless communication system that owns the exclusive rights to use the spectrum band of interest. In a bid to improve the spectrum usage efficiency while earning extra revenue, the primary system is willing to allow a secondary system to access this spectrum band whenever it can tolerate and to the maximum possible extent. It is further assumed that there are $K_p$ primary users in the primary system and there are $K_s$ secondary links of interest. For simplicity of exposition, all these secondary links are assumed to belong to the same secondary system. We will refer to $j$-th transmitter or $j$-th receiver to mean the transmitter and receiver of the $j$-th link. The channel gain between the $j$-th transmitter and the $k$-th receiver, either primary or secondary, is denoted by $h_{jk}$. We use $p_j$ to represent transmission power of the $j$-th user. Note that, depending on the type of wireless networks assumed, the receivers of each link may or may not be physically distinct. For simplicity we will assume that all primary users communicate with the same primary receiver (for example, a base station) although this assumption can easily be dropped at the expense of notational complexity.

In a C-DSL network, the primary system is assumed to adapt its interference cap (IC), denoted by $Q_0$, which is the maximum interference the *primary system* is willing to tolerate from all secondary transmissions at a given time, and thus its reward can be an increasing

function of the interference cap. However, in reality, the primary user should maintain a target signal-to-interference-plus-noise ratio (SINR) to ensure its required QoS. Moreover, an unnecessarily large interference cap by the primary user could hinder the performance of both systems due to resulting high primary interference. The goal of the secondary system, on the other hand, is to fully utilize the spectrum activity allowed by the primary user. Each secondary user may be assumed to act in its own interest to maximize its own utility. However, their transmission powers must be carefully self-regulated in order to ensure low interference to the primary user (within the IC) as well as to other secondary users.

The signal received at the primary receiver can be written as $r^{(p)}(t) = \sum_{i\in\mathcal{K}_p} A_{p,i}b_i s_i(t) + \sum_{j\in\mathcal{K}_s} A_{p,j}b_j s_j(t) + \sigma_p n(t)$, where $A_{p,l} = h_{pl}\sqrt{p_l}$ for $l \in \mathcal{K}_p \bigcup \mathcal{K}_s$, $n(t)$ is AWGN with unit spectral height and $\sigma_p^2$ is the variance of the zero-mean, additive noise at the primary receiver. Assuming $M$ discrete-time projections $r_m^{(p)} = \langle r^{(p)}(t), \psi_m^{(p)}(t)\rangle$, for $m = 1, \cdots, M$, of the continuous time signal on to a set of $M$ orthonormal directions specified by $\left\{\psi_1^{(p)}(t), \cdots, \psi_M^{(p)}(t)\right\}$, called the primary basis, and letting $\mathbf{r}^{(p)} = \left(r_1^{(p)}, \cdots, r_M^{(p)}\right)^T$, we may obtain the following discrete-time representation of the received signal at the primary receiver: $\mathbf{r}^{(p)} = \sum_{i\in\mathcal{K}_p} A_{p,i}b_i\mathbf{s}_i^{(p)} + \sum_{j\in\mathcal{K}_s} A_{p,j}b_j\mathbf{s}_j^{(p)} + \sigma_p\mathbf{n}^{(p)}$, where $\mathbf{s}_k^{(p)} = \left(s_{k1}^{(p)}, \cdots, s_{kM}^{(p)}\right)$, for $k \in \mathcal{K}_p$, is the $M$-vector representation of the $k$-th secondary user signalling waveform $s_k(t)$ w.r.t. the $M$-dimensional basis employed by the primary system, where $s_{km}^{(p)} = \langle s_k(t), \psi_m^{(p)}(t)\rangle$, and $\mathbf{n}^{(p)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_M)$. With the conventional matched-filter (MF) detector at the primary receiver, and assuming that primary modulation is BPSK so that $b_i \in \{+1, -1\}$, the $i$-th primary user symbols are detected as $\hat{b}_i = sgn\left(y_i^{(p)}\right)$ where $y_i^{(p)} = \left(\mathbf{s}_i^{(p)}\right)^T \mathbf{r}^{(p)} = A_{p,i}b_i + \sum_{l\in\mathcal{K}_p\setminus\{i\}} \rho_{il}^{(p)} A_{p,l}b_l + \sum_{j\in\mathcal{K}_s} \rho_{ij}^{(p)} A_{p,j}b_j + \sigma_p\eta_i^{(p)}$, with $\rho_{kl}^{(p)} = \left(\mathbf{s}_k^{(p)}\right)^T \mathbf{s}_l^{(p)}$ for $k, l \in \mathcal{K}_s \bigcup \mathcal{K}_s$ and $\eta_i^{(p)} \sim \mathcal{N}(0, 1)$. It is straightforward to observe that the total secondary interference (SI) from all secondary transmissions to the $i$-th primary user decisions is given by

$$I_i = \sum_{j\in\mathcal{K}_s} \left(\rho_{ij}^{(p)}\right)^2 h_{pj}^2 p_j = \sum_{j\in\mathcal{K}_s} \tilde{A}_{i,j}^2 p_j, \tag{9.1}$$

where $\tilde{A}_{i,j} = \rho_{ij}^{(p)} h_{pj}$. We denote the maximum of these interference at any given time over

all primary users as $I_0$, so that $I_0 = \max_{i\in\mathcal{K}_p} I_i = \sum_{j\in\mathcal{K}_s} \tilde{A}_j^2 p_j$ where $\tilde{A}_j = \tilde{A}_{i*j}$, for some $i* \in \mathcal{K}_p$. This total interference parameter $I_0$ plays a key role in the C-DSL based DSS systems, as we will see below.

Similarly, the received signal at the $j$-th secondary-system receiver can be written as $r_j^{(s)}(t) = \sum_{k\in\mathcal{K}_s} B_{j,k} b_k s_k(t) + \sum_{i\in\mathcal{K}_p} B_{j,i} b_i s_i(t) + \sigma_s n_j(t)$, where $B_{j,k} = h_{jk}\sqrt{p_k}$, for $k \in \mathcal{K}_s$, $B_{j,i} = h_{ji}\sqrt{p_i}$, for $i \in \mathcal{K}_p$ and $\sigma_s^2$ is the variance of secondary receiver noise. A discrete-time representation of $r_j^{(s)}(t)$ with respect to an $N$-dimensional orthonormal basis $\left\{\psi_1^{(s)}(t), \cdots, \psi_N^{(s)}(t)\right\}$ used by the secondary system, termed the secondary basis, can be written as $\mathbf{r}_j^{(s)} = \sum_{k\in\mathcal{K}_s} B_{j,k} b_k \mathbf{s}_k^{(s)} + \sum_{i\in\mathcal{K}_p} B_{j,i} b_i \mathbf{s}_i^{(s)} + \sigma_s \mathbf{n}_j^{(s)}$, where $\mathbf{r}_j^{(s)} = \left(r_{j1}^{(s)}, \cdots, r_{jN}^{(s)}\right)^T$, $r_{jn}^{(s)} = \langle r_j(t), \psi_n^{(s)}(t)\rangle$, for $n = 1, \cdots, N$, is the projection of the received signal at the secondary receiver $j$ on the the $n$-th orthonormal basis function, $\mathbf{s}_l^{(s)} = \left(s_{l1}^{(s)}, \cdots, s_{lN}^{(s)}\right)$, for $l \in \mathcal{K}_p \bigcup \mathcal{K}_s$, is the $N$-vector representation of $s_l(t)$ with respect to the $N$-dimensional basis employed by the secondary system with $s_{ln}^{(s)} = \langle s_l(t), \psi_n^{(s)}(t)\rangle$, and $\mathbf{n}_j^{(s)} \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_N)$.

### 9.1.3 C-DSL Game Model for Dynamic Spectrum Sharing

In the proposed C-DSL-based DSS networks, the primary system and secondary users interact with each other by adjusting their interference cap and transmit power levels, respectively, in order to maximize their own gains. The primary system action is to set the interference cap $Q_0$ which specifies the maximum interference it is willing to tolerate from all secondary users. We model the above system as in the following noncooperative C-DSL game $(\mathcal{K}, \mathcal{A}_k, u_k(.))$:

1. Players: $\mathcal{K} = \{0, 1, 2, \cdots, K_s\}$, where we assume that the 0-th user is the primary system consisting of multiple primary users and $k \in \mathcal{K}_s$ represents the $k$-th secondary user.

2. Action space: $\mathcal{P} = \mathcal{A}_0 \times \mathcal{A}_1 \times \mathcal{A}_2 \cdots \times \mathcal{A}_{K_s}$, where $\mathcal{A}_0 = \mathcal{Q} = [0, \bar{Q}_0]$ represents the

primary system's action set and $\mathcal{A}_k = \mathcal{P}_k = [0, \bar{P}_k]$, for $k \in \mathcal{K}_s$, represents the $k$-th secondary user's action set. Note that $\bar{Q}_0$ and $\bar{P}_k$ represent, respectively, the maximum IC of the primary system and the maximum transmission power of the $k$-th secondary user. We denote the action vector of all players in the noncooperative C-DSL game by $\mathbf{a} = [Q_0, p_1, \cdots, p_{K_s}]^T$, where $Q_0 \in \mathcal{Q}$ and $p_k \in \mathcal{P}_k$. The action vector excluding that of the $k$-th user is denoted as $\mathbf{a}_{-k}$.

3. Utility function: We denote by $u_0(Q_0, \mathbf{a}_{-0})$ the primary system's utility function and by $u_k(p_k, \mathbf{a}_{-k})$, for $k \in \mathcal{K}_s$, the $k$-th secondary user's utility function.

At any given time the target SINR of the $i$-th primary link is defined in terms of its assumed worst-case secondary interference $\bar{\gamma}_i = \frac{h_{pi}^2 p_i}{Q_0 + MAI(i) + \sigma_p^2}$, where $MAI(i) = \sum_{l \in \mathcal{K}_p \setminus \{i\}} \left(\rho_{il}^{(p)}\right)^2 h_{pl}^2 p_l = \sum_{l \in \mathcal{K}_p \setminus \{i\}} \left(\rho_{il}^{(p)}\right)^2 A_{p,l}^2$ is the *multiple access interference* (MAI) from all other primary transmissions to the $i$-th primary-user. Our proposed model allows primary users to adapt their actions so as to control their throughput. We could allow $\bar{\gamma}_i$ to be time-varying. In that case $Q_0(t)$ would have to be chosen in such a way so that $\gamma_i(t) \geq \bar{\gamma}_i(t)$ and secondary interference $I_0(t)$ would change according to $I_0(t) \leq Q_0(t)$. On the other hand, the $i$-th primary user's actual instantaneous SINR is given by

$$\begin{aligned}\gamma_i &= \frac{h_{pi}^2 p_i}{\sum_{l \in \mathcal{K}_p \setminus \{i\}} \left(\rho_{il}^{(p)}\right)^2 h_{pl}^2 p_l + \sum_{j \in \mathcal{K}_s} \left(\rho_{ij}^{(p)}\right)^2 h_{pj}^2 p_j + \sigma_p^2} \\ &\geq \bar{\gamma}_i \left(1 + \frac{Q_0 - I_0}{I_0 + MAI(i) + \sigma_p^2}\right). \end{aligned} \tag{9.2}$$

Thus, as seen from (9.2), each primary user's instantaneous SINR will be above the least acceptable SINR threshold as long as the primary system's interference cap $Q_0 \geq I_0$. It is to be noted that each primary user under the primary system choose its transmit power from least acceptable SINR threshold as $p_i = \bar{\gamma}_i \left(\frac{Q_0 + MAI(i) + \sigma_p^2}{h_{pi}^2}\right)$. Since $I_0 \leq Q_0$, instantaneous SINR of $i$-th primary user would be, in general, greater than the least acceptable SINR as seen from (9.2). If sharing is not enabled, i.e., $Q_0 = 0$ and $I_0 = 0$, $i$-th primary user's power would be $p_i = \bar{\gamma}_i \left(\frac{MAI(i) + \sigma_p^2}{h_{pi}^2}\right)$, which is less than the power, the $i$-th primary user had

to transmit if sharing were enabled. According to (9.2), the instantaneous SINR would be exactly equal to the least acceptable SINR in this case. On the other hand, if sharing is enabled, i.e. $Q_0 \neq 0$ and $I_0 \neq 0$, rate achieved by $i$-th primary user is $W_i \log(1+\gamma_i)$, which is at least greater than least acceptable data rate $W_i \log(1+\bar{\gamma}_i)$. Hence, if sharing is enabled, as long as $I_0 \leq Q_0$, data rate of $i$-th primary user is guaranteed to be above the minimum required threshold, but of course at the expense of transmitting at a higher power.

By generalizing the approach proposed in [96,99], we propose the following utility function for the primary system:

$$\begin{aligned} u_0\left(Q_0, \mathbf{a}_{-0}\right) &= u_0\left(Q_0, I_0\right) \\ &= \left(\bar{Q}_0 - \left(Q_0 - I_0(\mathbf{a}_{-0})\right)\right) F(Q_0), \end{aligned} \tag{9.3}$$

where $F(.)$ is a suitable continuous reward function for the primary system. For example, in [99] the authors proposed a linear reward function $F(Q_0) = Q_0$ assuming that the reward for the primary system is directly proportional to the interference cap it chooses. In this chapter, we establish conditions on $F(.)$ so that the proposed C-DSL game has desired equilibrium properties. Note that, (9.3) also assumes that the utility of the primary system is proportional to the *demand* in addition to the reward function $F(.)$. The demand is taken to be decreasing when extra interference margin $Q_0 - I_0$ increases. This discourages the primary system from swamping all other transmissions by setting too large an interference cap that will lead to higher transmission powers according to (9.2). As a special case of (9.3), we choose $F(Q_0) = \log(1 + Q_0)$ so that the primary utility is proportional to the capacity attained by the secondary system with respect to the primary receiver. We believe this model for the primary system utility is more sensible in a dynamic spectrum leasing cognitive radio network, compared to [99], when the secondary system is concerned about the rate its users achieve rather than their transmission powers. By choosing a reasonable revenue/utility rate based on the market value, the revenue earned by the primary system could be increased, and the revenue achieved in that case would reflect actual capacity achieved by the secondary system more compared to that in [99]. Figure 9.1 shows the above primary utility as a

function of the interference cap $Q_0$ for a fixed total secondary interference $I_0$. Observe from Fig. 9.1 that the primary system utility $u_0$ is quasi-concave in interference cap $Q_0$.

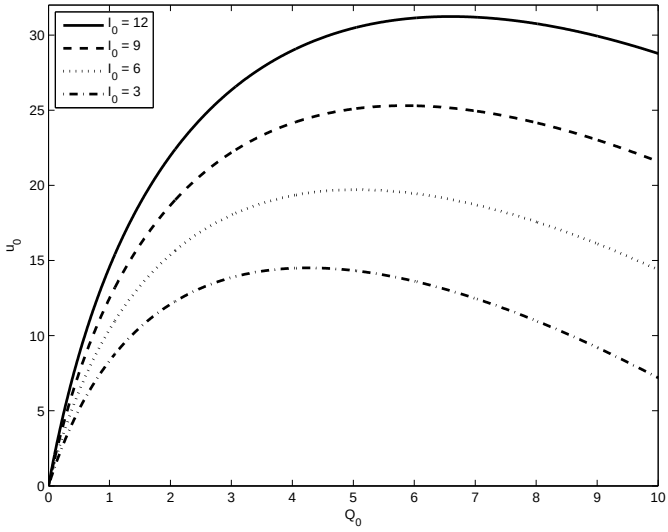



Figure 9.1: Primary utility $u_0$ for a fixed secondary interference $I_0$ with $\overline{Q}_0 = Q_0^{max} = 10$, $h_{p1} = 1$, $\rho_{01}^{(p)} = \rho_{10}^{(s)} = 1$ and $\lambda = 1$.

As can be seen from (9.2) as long as the secondary user interference $I_0 \leq Q_0$, the primary system quality of service will be guaranteed for all its users. To ensure this the utilities of secondary users must be fast decaying functions of $I_0 - Q_0$ when this difference is positive. Thus, if $f(p_k)$ is the reward achieved by the $k$-th secondary user by transmitting at a power $p_k$, then a suitable utility function for it would be $f(p_k)q(Q_0 - I_0)$ where $q(.)$ is the unit step function. Motivated by these arguments, and in an attempt to avoid the discontinuity of the step function, in this chapter we propose the following utility function for the $k$-th secondary user, for $k \in \mathcal{K}_s$:

$$\begin{aligned} u_k\left(p_k, \mathbf{a}_{-k}\right) &= \frac{f(p_k)}{1 + e^{\lambda(I_0 - Q_0)}} \\ &= \frac{f(p_k)}{1 + e^{\lambda\left(I_{0,-k}(\mathbf{a}_{-k}) - Q_0\right)} e^{c_k p_k}}, \end{aligned} \tag{9.4}$$

where $f(.)$ is a suitable reward function chosen by the secondary system and the weighting term $\frac{1}{1+e^{\lambda(I_0 - Q_0)}}$ in (9.4) is a sigmoidal function used to approximate the unit step with the

property that it goes to either +1 or 0, as $I_0 - Q_0$ tends to either negative or positive infinity, respectively, while the parameter $\lambda$ can be used to adjust the steepness of the transition region. Note that $I_{0,-k} = I_0 - \tilde{A}_k^2 p_k$ is the interference from all secondary transmissions to the worst-hit primary user excluding that from the $k$-th secondary user and that we have defined $c_k = \lambda \tilde{A}_k^2 \geq 0$ in (9.4).

In a dynamic spectrum leasing system a suitable objective for the secondary system would be to maximize the sum capacity of all its users in the shared primary spectrum. However, from the perspective of a particular secondary user, it would be interested in gaining the maximum possible rate it can achieve. As a result, we will use $f(p_k) = W_k \log\left(1 + \gamma_k^{(s)}\right)$ throughout this chapter as a special case of (9.4) where $\gamma_k^{(s)}$ is the received SINR of the $k$-th *secondary link* for $k \in \mathcal{K}_s$ and $W_k > 0$ can be taken as proportional to the bandwidth. Hence, in the remainder of this chapter we will limit ourselves to investigating equilibrium strategies of the game $G = (\mathcal{K}, \mathcal{A}_k, u_k)$ where users are interested in maximizing the utility functions defined in (9.3) and (9.4) with $F(Q_0) = \log(1 + Q_0)$ and $f(p_k) = W_k \log\left(1 + \gamma_k^{(s)}\right)$ respectively.

**Definition 1** *A strategy vector* $\mathbf{p} = (a_0, a_1, \cdots, a_k)$ *is a Nash equilibrium of the primary-secondary user power control game* $G = (\mathcal{K}, \mathcal{A}_k, u_k)$ *if, for every* $k \in \mathcal{K}$, $u_k(a_k, \mathbf{a}_{-k}) \geq u_k(a'_k, \mathbf{a}_{-k})$ *for all* $a'_k \in \mathcal{A}_k$.

The best response correspondence of a user in a game is the best reaction strategy a rational user would choose in order to maximize its own utility, in response to the actions chosen by other players.

**Definition 2** *The user k's best response* $r_k : \mathcal{A}_{-k} \longrightarrow \mathcal{A}_k$ *is the set*

$$r_k(\mathbf{a}_{-k}) = \{a_k \in \mathcal{A}_k : u_k(a_k, \mathbf{a}_{-k}) \geq u_k(a'_k, \mathbf{a}_{-k}) \quad \textit{for all} \;\; a'_k \in \mathcal{A}_k\}. \tag{9.5}$$

Clearly both the primary system and secondary user action sets are both compact and convex being closed and bounded intervals on the real line. Further, according to our construction, both $u_0(\mathbf{a})$ and $u_k(\mathbf{a})$ are continuous in the action vector $\mathbf{a}$. The usefulness of the best-response strategies come handy in establishing the uniqueness of the Nash equilibrium of a C-DSL game, as we will see later. Indeed it has been shown that if the best response correspondences $r_k(\mathbf{a}_{-k})$ of a game are so-called *standard functions* for every $k \in \mathcal{K}$, then the game has a unique Nash equilibrium [102], where

**Definition 3** *A function* $\mathbf{r}(\mathbf{a})$ *is said to be a standard function if it satisfies the following three properties [102]: (i) Positivity :* $\mathbf{r}(\mathbf{a}) > 0$*, (ii) Monotonicity : If* $\mathbf{a} \geq \mathbf{a}'$*, then* $\mathbf{r}(\mathbf{a}) \geq \mathbf{r}(\mathbf{a}')$*, (iii) Scalability : For all* $\mu > 1$*,* $\mu\mathbf{r}(\mathbf{a}) \geq \mathbf{r}(\mu\mathbf{a})$*.*

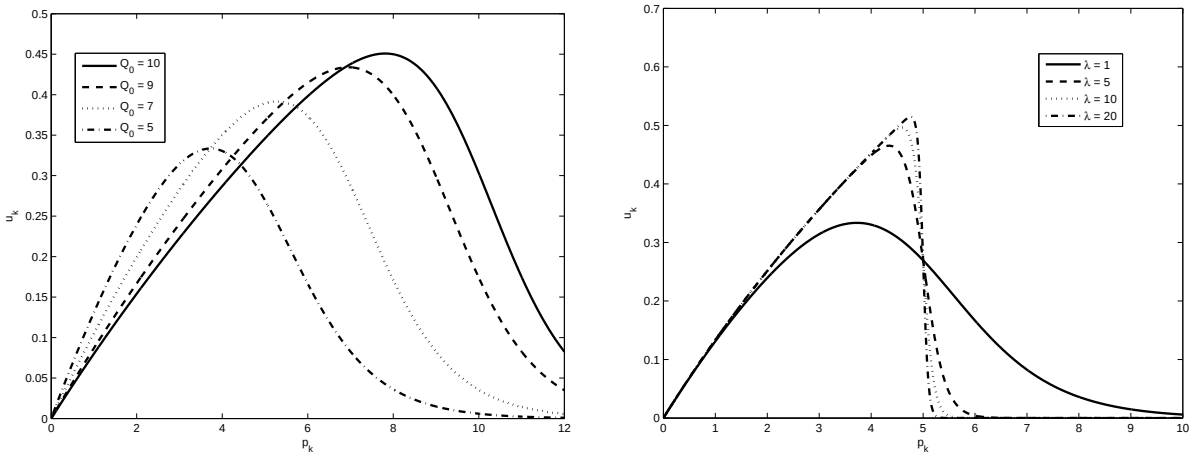
### 9.1.4 Analysis of The Proposed C-DSL Game with The MF Secondary Receiver

We assume that all secondary transmissions are BPSK and all secondary detectors are based on the MF. Then, the $j$-th secondary link receiver detects the corresponding $j$-th secondary transmitter's symbols as $\hat{b}_j = sgn\left(y_j^{(s)}\right)$ where, for $j \in \mathcal{K}_s$, $y_j^{(s)} = \left(\mathbf{s}_j^{(s)}\right)^T \mathbf{r}_j^{(s)} = B_{j,j}b_j + \sum_{l\in\mathcal{K}_s\ \{j\}} \rho_{jl}^{(s)} B_{j,l}b_l + \sum_{i\in\mathcal{K}_p} \rho_{ji}^{(s)} B_{j,i}b_i + \sigma_s\eta_j^{(s,j)}$ with $\rho_{kl}^{(s)} = \left(\mathbf{s}_k^{(s)}\right)^T \mathbf{s}_l^{(s)}$ and $\eta_k^{(s,j)} = \left(\mathbf{s}_k^{(s)}\right)^T \mathbf{n}_j^{(s)} \sim \mathcal{N}(0,1)$. Hence, the $j$-th secondary link's SINR is given by

$$\begin{aligned}\gamma_j^{MF} &= \frac{|h_{jj}|^2 p_j}{\sum_{l\in\mathcal{K}_s\setminus\{j\}} \left(\rho_{jl}^{(s)}\right)^2 |h_{jl}|^2 p_l + \sum_{i\in\mathcal{K}_p} \left(\rho_{ji}^{(s)}\right)^2 |h_{ji}|^2 p_i + \sigma_s^2} \\ &= \frac{|h_{jj}|^2 p_j}{i_j^{(j)} + \tilde{\sigma}_{s,j}^2} \;=\; \frac{p_j}{N_j}, \end{aligned} \tag{9.6}$$

where $i_k^{(j)} = \sum_{l\in\mathcal{K}_s\setminus\{k\}} \left(\rho_{kl}^{(s)}\right)^2 h_{jl}^2 p_l$ is the total interference from all other secondary users to the $k$-th secondary link signal at the $j$-th secondary receiver, $\tilde{\sigma}_{s,j}^2 = \sum_{i\in\mathcal{K}_p} \left(\rho_{ji}^{(s)}\right)^2 |h_{ji}|^2 p_i + \sigma_s^2$

is the effective primary interference plus noise seen by the $j$-th link and $N_j = \frac{i_j^{(j)} + \tilde{\sigma}_{s,j}^2}{|h_{jj}|^2}$. It can be easily seen that $\frac{\partial \gamma_j^{MF}}{\partial p_j} = \frac{\gamma_j^{MF}}{p_j}$ when the secondary system employs the MF receiver.




(a) (b)

Figure 9.2: Secondary-link utility $u_k$ for (a) fixed interference cap $Q_0$ and (b) fixed $\lambda$ with $Q_0 = 5$. Other parameters used are: $W_k = W = 1$, $h_{p1}^2 = h_{11}^2 = 1$, $\sigma_p^2 = \sigma_s^2 = 1$ and all the cross-correlations are assumed to be unity.

Assuming single-user primary and secondary systems, the proposed secondary user utility in (9.4) with $\gamma_j^{(s)} = \gamma_j^{MF}$ is shown in Fig. 9.2(a) parameterized by the primary interference cap, while Fig. 9.2(b) shows the effect of parameter $\lambda$ on the utility function. It can be seen from Fig. 9.2(a) that the proposed secondary user utility function $u_j$ is quasi-concave in $p_j$, and the unique maximum of $u_j$ is an increasing function of primary interference cap $Q_0$. Hence pushing the primary interference cap to a higher value encourages the secondary users to transmit at higher powers and thus allow the primary user to achieve higher leasing gains. From Fig. 9.2(b), it can be seen that the parameter $\lambda$ can be used to adjust the steepness of the transition region of the secondary utility function. The higher the value of $\lambda$, the steeper the transition, and indicates that the primary system expects the secondary users to strictly adhere to the $I_0 \leq Q_0$ requirement. The following proposition guarantees the existence of a Nash equilibrium in the proposed centralized DSL game under certain conditions on the

primary reward function $F(Q_0)$:

**Proposition 1** *With $\mathcal{A}_k$'s and $u_k$'s as defined above, the dynamic spectrum leasing (DSL) game $G = (\mathcal{K}, \mathcal{A}_k, u_k)$ has a Nash Equilibrium when $\gamma_k^{(s)} = \gamma_k^{MF}$ if $F(Q_0)$ satisfies the following conditions:*

1. *$F(Q_0)$ is continuous and strictly monotonic for $Q_0 > 0$*
2. *$F(0) = 0, F'(0) > 0$ and $\lim_{Q_0 \to \infty} \frac{F(Q_0)}{F'(Q_0)} > -\infty$*
3. *$\frac{F(Q_0)F''(Q_0)}{(F'(Q_0))^2} < 2$ for $Q_0 > 0$*
4. *$0 \leq \bar{Q}_0 + I_0(\mathbf{a}_{-0}) < \infty$.*

*Proof: See Appendix 9A.* ■

In a non-cooperative game if all users are allowed to adapt their actions sequentially according to their best-response correspondences, then they are guaranteed to converge to a Nash equilibrium of the game. The unique interior best response of primary system is given by the solution to $Q_0^*(I_0) = \left(\bar{Q}_0 + I_0\right) - \frac{F(Q_0^*)}{F'(Q_0^*)}$. Since $u_0(Q_0)$ is monotonic increasing for $Q_0 < Q_0^*$, if the maximum interference cap is such that $\bar{Q}_0 < Q_0^*$, then the primary system best response is given by $r_0(\mathbf{a}_{-0}) = \min\{\bar{Q}_0, Q_0^*(I_0)\}$. In order to determine this best response $r_0(\mathbf{a}_{-0})$ for a chosen power vector by the secondary links, the only quantity the primary system needs to know is the maximum total secondary interference experienced by any user at the primary receiver, denoted by $I_0 = \max_{i \in \mathcal{K}_p} I_i$. This total interference can easily be estimated at the primary receiver without much difficulty. It is to be noted that the simplification $I_0 = \max_i I_i$ may lead to not fully capitalizing on different channel conditions on different primary users, as different primary users may experience different interference from secondary users in general[1]. However, the context to which the proposed

[1]Exploiting the channel variations over different primary channels/frequency bands is considered in our follow-up paper [103].

C-DSL applies is when a set of primary users share the same frequency band (e.g. CDMA) and communicate to a single receiver. In that case, the receiver needs to be able to work with the worst-hit primary user conditions.

On the other hand, the best response of the $j$-th secondary link to the transmit powers of other secondary users as well as IC set by the primary user is given by the (unique) solution $p_j = p_j^*\left(Q_0, I_{0,-j}, i_j^{(j)}\right)$ to the equation $g_j^{MF}(p_j) = 0$ defined in Appendix 9.1.7. Again, since $u_j$ is quasi-concave in $p_j$, if $p_j^*\left(Q_0, I_{0,-j}, i_j^{(j)}\right) > \bar{P}_j$, then its best response is to set its transmit power to $p_j = \bar{P}_j$. Hence, we have the best response of the $j$-th secondary link, for $j \in \mathcal{K}_s$: $r_j\left(\mathbf{a}_{-j}\right) = \min\{\bar{P}_j, p_j^*\left(Q_0, I_{0,-j}, i_j^{(j)}\right)\}$.

Observe that in general the best response of the $j$-th secondary link is a function of the residual interference $Q_0 - I_{0,-j}$ of all other secondary users at the primary receiver and the total interference from all secondary and primary users to the $j$-th link at its receiver. The $j$-th secondary link receiver can easily estimate the latter quantity. However, to obtain the residual interference $Q_0 - I_{0,-j}$ the secondary receiver needs to know the current interference cap $Q_0$ as well as the secondary interference $I_{0,-j}$ at the primary receiver. In this work, we assume that the primary base station broadcasts both $Q_0$ and $I_0$ whenever it adjusts its interference cap to a new value. This is the only conscious interaction the primary system is assumed to be having with the secondary system if decentralized optimization is considered. Observe that knowing $I_0$, each secondary user can compute the residual interference $I_{0,-j} = I_0 - \tilde{A}_j^2 p_j$ since it knows its own transmit power and it may estimate the channel state information $\tilde{A}_j$ if the reverse link signals are available on the same frequency band. On the other hand, if the optimization were to be centralized, then the central system had to be aware of the channel state information and powers of all the secondary users. In addition to that it had to inform each individual secondary user about its new transmit power $p_k$, which would cost sufficient amount of bandwidth dedicated to control signals.

### 9.1.5 Analysis of The Proposed C-DSL Game with the LMMSE Secondary Receiver

In this section we assume that the secondary system is equipped with so-called LMMSE receivers. Note that, to make fair comparisons with the case of MF-based secondary receivers as discussed in the previous section, we hold the primary receiver to be still based on the MF. Of course it is also possible for the primary system to be equipped with an LMMSE (or any other MUD) detector. The effect of that would be for the primary system to be able to tolerate higher $I_0$ values, for the same primary transmission powers.

For detecting signals of the $j$-th secondary link, the $j$-th secondary link receiver employs the LMMSE filter defined by $\min_{\mathbf{w_{j,j}}} \mathbb{E}\left[\left(b_j - \mathbf{w}_{j,j}\mathbf{r}_j^{(s)}\right)^2\right]$, where $\mathbf{w}_{j,j} \in \mathbb{R}^N$ is the vector of LMMSE filter coefficients at the $j$-th receiver that achieves minimum mean-squared error in estimating $j$-th link symbols. It is well-known that the solution to the above optimization problem is given straightforwardly by $\mathbf{w}_{j,j} = \mathbb{E}\left[\mathbf{r}_j^{(s)}\left(\mathbf{r}_j^{(s)}\right)^T\right]^{-1}\mathbb{E}\left[b_j\mathbf{r}_j^{(s)}\right]$. It can be verified that $\mathbb{E}\left[\mathbf{r}_j^{(s)}\left(\mathbf{r}_j^{(s)}\right)^T\right] = \sum_{l\in\mathcal{K}_s} B_{j,l}^2\mathbf{s}_l^{(s)}\mathbf{s}_l^{(s)^T} + \sum_{i\in\mathcal{K}_p} B_{j,i}^2\mathbf{s}_i^{(s)}\mathbf{s}_i^{(s)^T} + \sigma_s^2\mathbf{I}$ and $\mathbb{E}\left[b_j\mathbf{r}_j^{(s)}\right] = B_{j,j}\mathbf{s}_j^{(s)}$, resulting in the following LMMSE filter coefficient vector for the $j$-th link: $\mathbf{w}_{j,j} = \frac{B_{j,j}}{1+B_{j,j}^2\mathbf{s}_j^{(s)^T}\Sigma_{j,j}^{-1}\mathbf{s}_j^{(s)}}\Sigma_{j,j}^{-1}\mathbf{s}_j^{(s)}$, where $\Sigma_{j,j} = \sigma^2\mathbf{I} + \sum_{i\in\mathcal{K}_p} B_{j,i}^2\mathbf{s}_i^{(s)}\mathbf{s}_i^{(s)^T} + \sum_{l\in\mathcal{K}_s\setminus\{j\}} B_{j,l}^2\mathbf{s}_l^{(s)}\mathbf{s}_l^{(s)^T}$. Note that although we omit details due to space constraints, the above LMMSE filter coefficient vector can easily be adapted without explicit knowledge of primary or the other secondary signaling waveforms. We refer the interested readers to [104]. The received output SINR of the $j$-th secondary link can be written as:

$$\begin{aligned}\gamma_j^{MMSE} &= B_{j,j}^2\left(\mathbf{s}_j^{(s)}\right)^T\Sigma_{j,j}^{-1}\mathbf{s}_k^{(s)} \\ &= h_{jj}^2 p_j\left(\mathbf{s}_j^{(s)}\right)^T\Sigma_{j,j}^{-1}\mathbf{s}_j^{(s)}. \end{aligned} \tag{9.7}$$

Due to the LMMSE detector's well known property of maximizing the output SINR, the linear MMSE receiver may require secondary radios to transmit at a lower power than that with the MF receiver to achieve the same QoS. Note that, since $\mathbf{s}_j^{(s)}$, $\Sigma_{j,j}^{-1}$ and $h_{jj}$ are

independent of $p_j$, it follows that $\frac{\partial \gamma_j^{MMSE}}{\partial p_j} = \frac{\gamma_j^{MMSE}}{p_j}$ as in the case of MF based receivers. In fact, it should be pointed out that the only difference between the C-DSL game with the LMMSE receiver and that with the MF receiver is in the above SINR expression of secondary users. As a result we have the following proposition, whose proof has been deferred to the Appendix 9.1.7, that establishes the existence and uniqueness properties of the equilibrium of the C-DSL game when secondary receivers are equipped with LMMSE detectors:

**Proposition 2** *With $\mathcal{A}_k$'s and $u_k$'s as defined before, the centralized dynamic spectrum leasing game $G = (\mathcal{K}, \mathcal{A}_k, u_k)$ still has a* unique *Nash Equilibrium when $\gamma_k^{(s)} = \gamma_k^{MMSE}$, if conditions $(1) - (4)$ in Proposition 1 are satisfied.*

*Proof: See Appendix 9B.* ■

### 9.1.6 Performance Analysis of a Centralized Dynamic Spectrum Leasing System

In this section, our goal is to investigate the behavior of the primary and secondary systems at the equilibrium and delineate the key characteristics emerging from our framework for spectrum leasing. We will compare performance of our proposed framework for both MF and LMMSE secondary receivers. The performance of the system is considered as its performance at the Nash equilibrium. For simplicity of exposition, we assume that both primary and secondary systems are equipped with only one receiver each in the uplink.

#### Identical links: AWGN Channels

To illustrate the characteristics of the Nash equilibrium in this primary-secondary user C-DSL game, it is interesting to look at perhaps the most simple situation in which there are identical secondary links ($K_s > 1$) and a single primary user ($K_p = 1$). We assume that all

secondary links' have the same cross-correlation coefficients $\rho_{0k}^{(p)} = \rho_0^{(p)}, \rho_{k0}^{(s)} = \rho_0^{(s)}, \rho_{kj}^{(s)} = \rho^{(s)}$, for all $k, j \in \mathcal{K}_s$ and all channels are additive white Gaussian noise (AWGN): $h_{sk} = h_{pk} = 1$ for all $k \in \mathcal{K}_s$ so that $\tilde{A}_k = \tilde{A}$ for all $k$. By symmetry, in this case all secondary users must have the same power $p_k = p^*$ at the Nash equilibrium (equivalently, the same SINR $\gamma_k = \gamma^*$). Thus when the secondary system employs MF receiver, with $F(Q_0) = \log(1 + Q_0)$, the Nash equilibrium is characterized by the intersection $(Q_0^*, p^*)$ of the following two curves:

$$Q_0 = r_0(p) = (\text{solution to equation } \psi_{Q_0}(p) = 0), \tag{9.8}$$

$$p = r_s(Q_0) = \big(\text{solution to equation } g^{MF}(p) = 0\big), \tag{9.9}$$

where $\psi_{Q_0}(p) = Q_0 + (1 + Q_0)\log(1 + Q_0) - \bar{Q}_0 - K\tilde{A}^2 p$, and $g^{MF}(p) = \frac{1}{N}e^{-cp} + \frac{1}{N}e^{-\lambda Q_0}e^{(K-1)cp} - c\left(1 + \frac{p}{N}\right)\log\left(1 + \frac{p}{N}\right)e^{-\lambda Q_0}e^{(K-1)cp}$. In the case of a secondary system with the LMMSE receiver, the Nash equilibrium is given by the intersection $(Q_0^*, p^*)$ of (9.8) and the curve: $p = r_s(Q_0) = \big(\text{solution to equation } g^{MMSE}(p) = 0\big)$, where $g^{MMSE}(.)$ is defined in Appendix 9.1.7 for identical secondary links.

Figure 9.3 shows the C-DSL game outcomes when the secondary system is equipped with the MF as well as the LMMSE receiver. Note that we have set $W_k = W = 1$, $\bar{Q}_0 = 10$, $\bar{P}_k = 20$, $\bar{\gamma}_0 = 1$, $h_{pk}^2 = h_{sk}^2 = 1$, for all $k \in \mathcal{K}_s$, $\sigma_s^2 = \sigma_p^2 = 1$, and all cross-correlations being 0.5. While the system shows similar performance trends with both receivers, the effect of having LMMSE receiver is that the safety margin $Q_0 - I_0$ is slightly larger compared to that with the MF receiver. As can be seen from Fig. 9.3, with the MF-based secondary receiver, the primary system can support only up to $K_s \leq 13$ secondary users before the secondary system violates the primary interference cap. On the other hand, with an LMMSE-based secondary receiver, the primary system can support up to $K_s \leq 15$ secondary users due to the superior interference suppression capability of the LMMSE receiver [104].

Figure 9.4 shows the primary and secondary utilities at the Nash equilibrium of the system as a function of the secondary system size $K_s$. It is observed from Fig. 9.4(a) that the

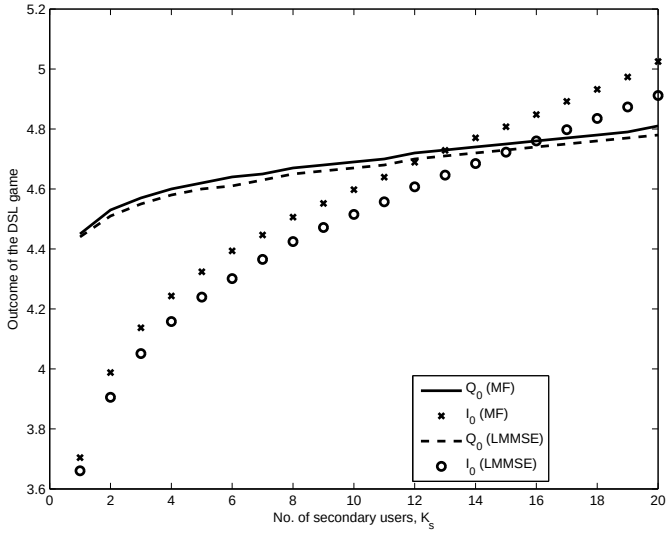
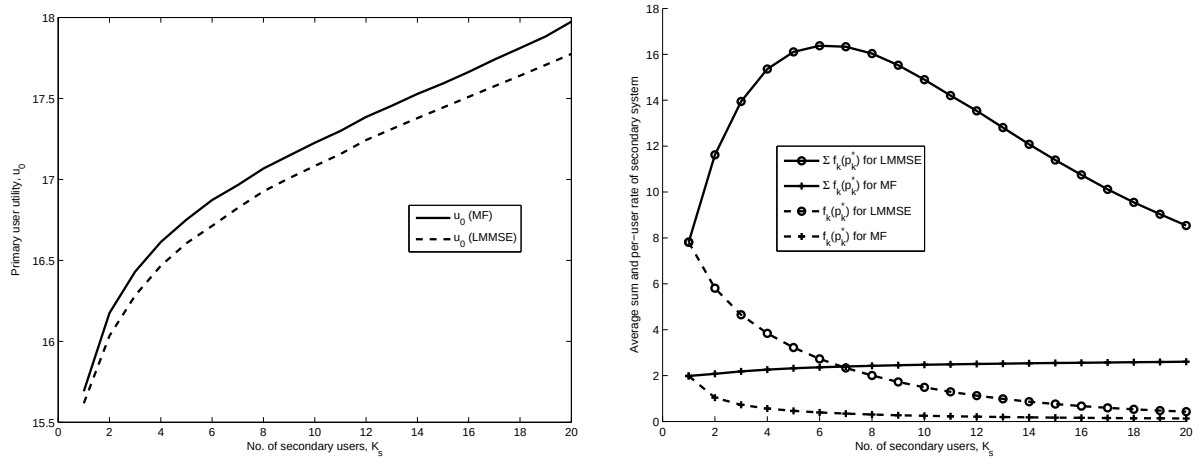



Figure 9.3: Outcome of the C-DSL game at the system NE, with MF and LMMSE receiver, as a function of secondary system size $K_s$ assuming identical secondary links, when $\lambda = 5$.




(a) (b)

Figure 9.4: Primary and secondary utilities at the system NE as a function of secondary system size $K_s$ for $\lambda = 5$ and assuming identical secondary user. (a) Primary system utility, (b) Sum-rate and the per-user rate achieved by the secondary system at the NE.

equilibrium utility of the primary system is decreased when the secondary system is equipped with the LMMSE receiver. This is because with the LMMSE receiver, the secondary system

can better manage its transmit power and thus total secondary interference to the primary system is reduced, which in turn reduces the primary utility. If one were to interpret the primary utility as proportional to a leasing payment the secondary system needs to make to the primary system, this shows how the secondary system can benefit by employing better receiver techniques.

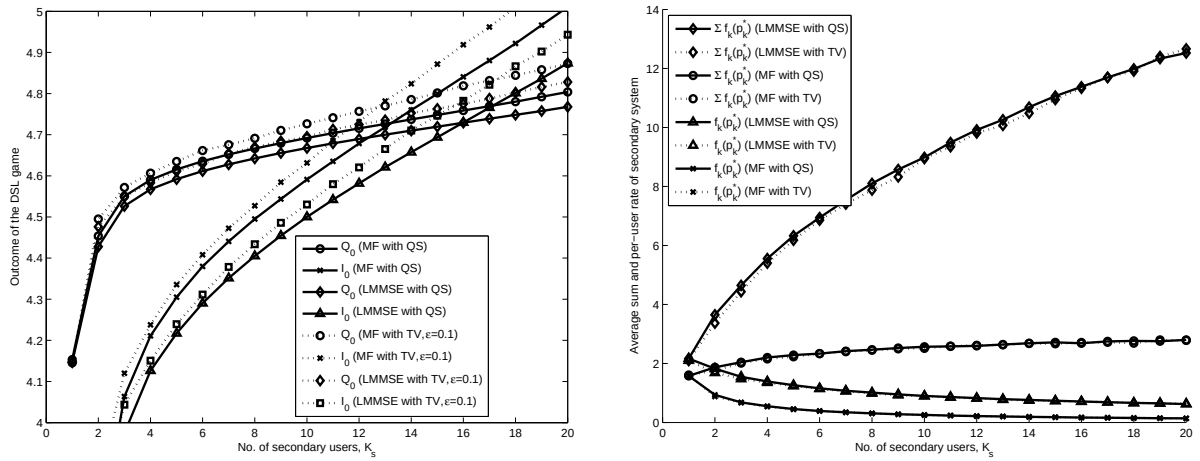
Figure 9.4(b) shows both the sum-rate $\sum_{k=1}^{K_s} \log\left(1+\gamma_k^{(s)}\right)$ as well as the per-user rate $\frac{1}{K_s}\sum_{k=1}^{K_s} \log\left(1+\gamma_k^{(s)}\right)$ achieved by the secondary system at the Nash equilibrium. As we observe from Fig. 9.4(b), the secondary sum-utility and per-user utility with the LMMSE receiver are higher compared to those achieved with the MF receiver. It can also be seen that the sum-utility of all the secondary users with LMMSE receiver has a unique maximum at $K_s = 6$. As the secondary system attempts to include more users into the system, the sum-utility of the secondary system starts to monotonically decrease. This is because, as the number of secondary users increases, users in both primary and secondary system experience additional interference. In response to that, each secondary user attempts to transmit at a higher power to achieve their target SINR's, thereby causing an overall degradation of both sum- and per-user rate. Note that, from a system point of view the secondary system would prefer to maximize the sum-rate. Thus from the secondary system's perspective, it may prefer to operate at $K_s = 6$ with the LMMSE receiver. As we see from Fig. 9.4(b), the sum-rate first increases and then decreases with $K$ for LMMSE-based receiver, but stays almost the same for MF-based receiver. Thus, at a first glance, allowing more secondary users to operate simultaneously seems to be the preferred solution with the MF-based receiver. However, Figure 9.4(b) also shows that the per-user rate is monotonically decreasing in $K_s$ for both MF and LMMSE-based secondary system, leading to decreasing incremental gains in sum-rate (with the MF receiver) as additional secondary users are added to the system. Depending on the application and the QoS requirement of the secondary system, each secondary user may have a minimum required rate (in bits per transmission) below which the transmissions would be useless. Thus we note that this QoS requirement will determine the maximum number of secondary users, the system would want to support at

any given time. For example, with the LMMSE-based secondary system, if the minimum per-user rate required is 2 bps, the optimal $K_s$ would be $K_s^* = 8$. If, on the other hand, the rate threshold was reduced to 1 bps, the secondary system might allow up to $K_s^* = 12$ secondary users to operate simultaneously. Note that on the other hand, if maximizing the sum-rate were to be the objective, then as noted above the optimum $K_s$ would be $K_s^* = 6$.

**Non-identical links: Fading Channels**

In the presence of wireless channel fading, the Nash equilibrium power profile of the C-DSL based dynamic spectrum sharing system will depend on the observed channel state realizations as well as on the type of receivers used in the secondary system. It is expected that in this case the Nash equilibrium transmit powers of individual secondary users will be different for each user. We assume that all channel gains follow Rayleigh distributions with all channel coefficients normalized so that $\mathbb{E}\{h^2\} = 1$. Other parameters used for simulations are: $W_k = W = 1$, $\bar{Q}_0 = 10$, $\bar{P}_k = 20$, $\lambda = 5$, $\bar{\gamma}_0 = 1$, $\sigma_p^2 = \sigma_s^2 = 1$ and all cross-correlations being 0.5. We investigate the performance with both quasi-static (QS) fading (channel state information is constant for the duration of a block) and slow time varying fading. For slow time varying (TV) fading, the temporal correlation is modeled as a first order Gauss-Markov process [105], and is described via $h_{(.,.)}(i) = \sqrt{1-\epsilon^2} h_{(.,.)}(i-1) + \epsilon w_{(.,.)}(i)$, where the driving noise $w_{(.,.)}(i)$ are iid $\mathcal{CN}(0, \sigma^2_{h_{(.,.)}})$ and $\epsilon$ is the channel variation rate. We assume that the channel state information (CSI) is not instantaneously available to the receivers, and each receiver updates the CSI periodically every $L$ samples. The detectors' decisions will use the estimated CSI defined as: $\hat{h}_{(.,.)}(i) = h_{(.,.)}(\lfloor i/L \rfloor L)$. For our simulations, we used $L = 10$ and $\epsilon = 0.1$. All simulation results are obtained by averaging over 2000 fading realizations.

In Fig. 9.5(a), we have shown the C-DSL game outcome at the Nash equilibrium as a function of number of secondary users $K_s$ in the presence of both time-varying and quasi-static channels. It can be seen from Fig. 9.5(a) that for quasi-static channel and with secondary MF receiver, the primary system can support up to $K_s \leq 13$ secondary users before




(a) (b)

Figure 9.5: (a) Outcome of the C-DSL game at the system NE, with MF and LMMSE secondary receiver, as a function of secondary system size $K_s$ in the presence of channel fading, (b) Average sum-rate and the per-user rate achieved by the secondary system at the NE.

the secondary system violates the primary interference cap. With the secondary LMMSE receiver, the primary system can support up to $K_s \leq 16$ secondary users in this channel scenario. On the other hand, for time-varying channel, the primary system can support up to $K_s \leq 12$ and $K_s \leq 15$ with the secondary MF and LMMSE receivers respectively. Note also that in the time-varying fading scenario, the safety margin $Q_0 - I_0$ decreases as secondary system size increases. In the time-varying case, values of $Q_0$ and $I_0$ are slightly higher than those for the quasi-static system. Figure 9.5(b) shows that the secondary link utility is decreased in time varying channels compared to that with quasi-static channels. This is because of the incomplete channel information forcing the system to deviate from the actual Nash equilibrium. It also shows that the secondary sum-utility as well as the per-user utility with the LMMSE receiver are also much better compared to those achieved with the MF receiver. Note that, the monotonic reduction in per-user utility with $K_s$ is common to both LMMSE and MF-based receivers. However, with the LMMSE receiver, this monotonic reduction is more than offset by the increased number of users in the secondary system.

When $f(p_k) = \log\left(1+\gamma_k^{(s)}\right)$ the reward for a secondary link is the rate (in bps) it can achieve assuming all other transmissions (both primary and secondary) are purely noise. The minimum transmission quality for the secondary system is defined as the average (over fading) minimum reward achieved by a link at the equilibrium. We denote this minimum required QoS for secondary link $k$ as $f_{min,k}$ and in all simulation results below assume that $f_{min,k} = f_{min}$ for all secondary links. Figure 9.6 shows the outage probability

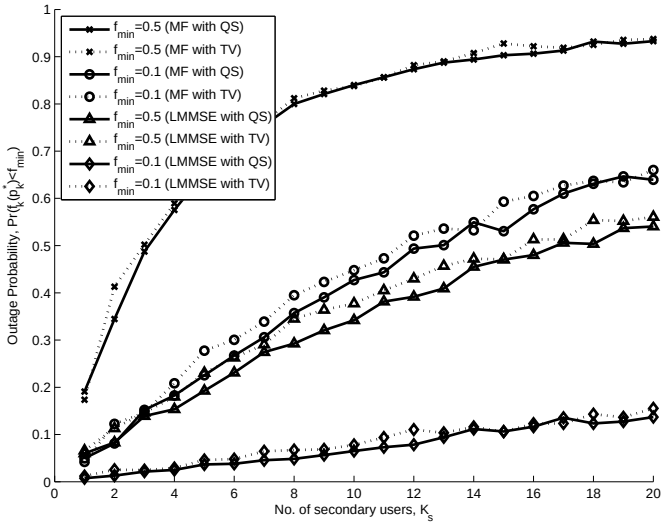
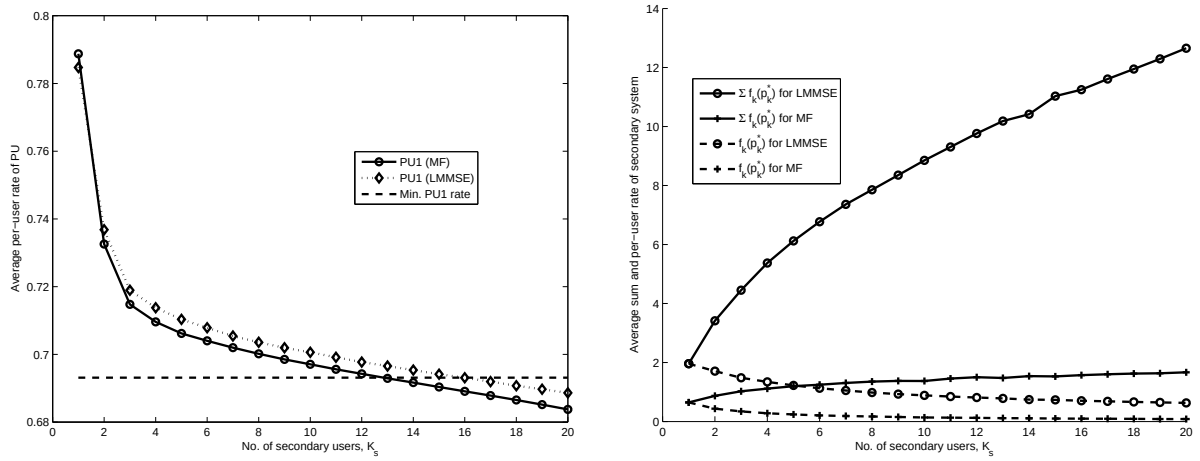



Figure 9.6: Outage probability $Pr\left(f_k\left(p_k^*\right) < f_{min}\right)$ of a typical secondary user at the NE of the C-DSL game in fading channels as a function of secondary system size $K_s$ for a required QoS requirement $f_{min}$.

$Pr\left(f_k\left(p_k^*\right) < f_{min}\right)$ of a typical secondary link as a function of $K_s$. It can be seen from Fig. 9.6 that the outage probability increases with $K_s$ as well as with the minimum QoS requirement. Note also that in general outage probability increases for both secondary MF and LMMSE receivers. Here also the LMMSE-based system ensures a higher QoS due to efficient management of secondary links' transmit powers. The maximum secondary system size that can be supported according to Fig. 9.5(b) thus needs to be interpreted in conjunction with the outage probabilities shown in Fig. 9.6. For example, in quasi-static channels with a MF-based secondary system, Fig. 9.5(b) shows that about 4 secondary links can

(on average) meet the $f_{min} = 0.5$ bps QoS requirement with $\lambda = 5$. However, according to Fig. 9.6 each of the these users may be in outage about 60% of time. On the other hand, for the LMMSE-based secondary system, about 20 secondary links can (on average) meet $f_{min} = 0.5$ bps QoS in Fig. 9.5(b), while also having an outage probability of about 50% according to Fig. 9.6. This, of course, is the price of operating as the secondary system. However, outage probabilities with the model proposed in this chapter are better than that in [99]. The improvement in the outage probability $f_{min} = 0.1$ bps with secondary MF receiver, is almost 8 times for $K_s = 1$ and 1.5 times for $K_s = 20$. On the other hand for the same $f_{min}$, improvement in the outage probability with the secondary LMMSE receiver, is almost 20 times for $K_s = 1$ and 7 times for $K_s = 20$. This is because of the difference in the secondary utility functions $u_k$'s in this chapter and in [99]. Since we approximate the unit step function with a sigmoidal function, with a suitably chosen value of the parameter $\lambda$ and higher value of positive $Q_0 - I_0$, the secondary utility function $u_k$ converges to the reward function $f(p_k)$. Thus unlike the secondary utility function $u_k$ in [99], the maximum of our proposed secondary utility function $u_k$ also approximately corresponds to maximum of $f(p_k)$. As a result, outage probability here is better than that in [99].

Figure 9.7 shows the primary and secondary rates achieved at the NE in a quasi-static channel as a function of secondary system size $K_s$ when the primary system has $K_p = 3$ users. It can be seen from Fig. 9.7(a), the primary user data rate is above the minimum required threshold as long as $I_0 \leq Q_0$ and it decreases as more secondary users are added into the system. For $K_s \geq 13$ (MF) and $K_s \geq 15$ (LMMSE), $I_0 \geq Q_0$ and the primary user rate drops below minimum required threshold. Hence, operation beyond this points is not desirable. As one would expect, in Fig. 9.7(b), the secondary sum and per-user rates are decreased with both the MF and LMMSE detector for $K_p = 3$. This is because of the increase in the total interference due to additional primary users in the system. As a result, the secondary system with the MF detector suffers more than that with the LMMSE detector when additional primary users are accommodated in the primary system. In Fig. 9.8, due to the same reason given above, the QoS of the secondary system in terms of outage worsens




(a) (b)

Figure 9.7: Primary and secondary rates at the system NE as a function of secondary system size $K_s$ in the presence of channel fading with $K_p = 3$ (a) Average data rate of primary user 1, (b) Average sum-rate and the per-user rate achieved by the secondary system at the NE.

with both the LMMSE and MF detectors for $K_p = 3$.

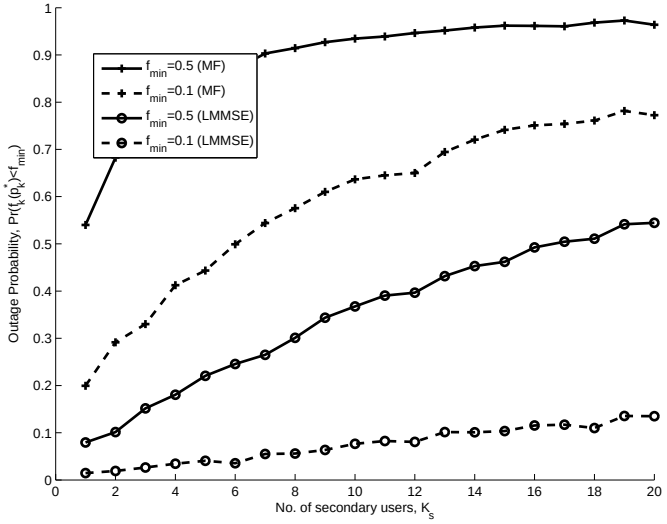



Figure 9.8: Outage probability $Pr\left(f_k\left(p_k^*\right) < f_{min}\right)$ of a typical secondary user at the NE of the C-DSL game in fading channels with $K_p = 3$ for a required QoS requirement $f_{min}$.

### 9.1.7 Conclusion

In this chapter, we developed a new game-theoretic framework for dynamic spectrum sharing in cognitive radio networks. In contrast to previously proposed hierarchical dynamic spectrum access networks, the proposed *centralized dynamic spectrum leasing* (C-DSL) networks provide an incentive for the primary system (possibly containing multiple primary users) to proactively accommodate secondary spectrum access whenever feasible. In our proposed framework, motivated by the network utility considerations, we further generalized the primary system utility defined in [96, 99] and have introduced a new utility function for the secondary system. We generalized our proposed game to allow for linear multiuser detectors, such as MF and LMMSE receivers at the secondary system. We established the conditions on the existence and uniqueness of the Nash equilibrium. In particular, we have established the general conditions on the primary system reward functions $F(.)$ so as to ensure the existence of a unique Nash equilibrium. We analyzed several examples of C-DSL networks in detail to investigate the proposed system behavior at equilibrium. For such a system, we observed that the proposed C-DSL game leads to a design that determines the maximum number of secondary links based on the required minimum QoS along with the $I_0 \leq Q_0$ criteria. We showed that the secondary system with the LMMSE receiver outperforms that with the MF receiver in terms of both the allowed secondary system size and the outage probability. It is to be noted that the meaning of *best performance* of the proposed C-DSL game may depend on what performance aspect we are interested in. From the secondary system point of view, C-DSL with LMMSE secondary receiver is better since it maximizes the output SINR. However, the primary system utility decreases when the secondary receiver is equipped with the LMMSE detector, which leads to loss in revenue. Hence, from primary perspective LMMSE secondary receiver may not be preferred. We also investigated, through simulations, the robustness of the proposed C-DSL game to slow time-varying fading, and showed that the primary and secondary systems can still successfully coexist at the C-DSL Nash equilibrium. Finally we showed that the proposed C-DSL game performs well even

when there are more than one primary users active in the spectrum band of interest.

# Appendix 9A

## Existence and Uniqueness of a NE with The Secondary MF Receiver

With the assumed form of action sets $\mathcal{A}_k$, the best response $r_k(\mathbf{a}_{-k})$ is both compact and convex for all $k \in \mathcal{K}_s$. Further, both $u_0(\mathbf{p})$ and $u_k(\mathbf{p})$ are continuous in the action vector $\mathbf{p}$. Let us define a function $\Phi(Q_0)$ as, $\Phi(Q_0) = \frac{F(Q_0)}{F'(Q_0)} + Q_0$. It can be seen that $u_0$ has a local maximum that is indeed a global maximum if $\Phi(Q_0) = \bar{Q}_0 + I_0(\mathbf{a}_{-0})$ has only one solution for $Q_0 \in \mathcal{Q}$. Clearly this equation has a solution if $\Phi(Q_0)$ is continuous and $\lim_{Q_0 \to 0} \Phi(Q_0) \leq \bar{Q}_0 + I_0(\mathbf{a}_{-0}) < \lim_{Q_0 \to \infty} \Phi(Q_0)$. This solution would be a global maximum if in addition $\Phi'(Q_0) > 0$ for $Q_0 > 0$. It can be easily verified that $\Phi'(Q_0) > 0$ will be true if $F(Q_0)$ is such that $\frac{F(Q_0)F''(Q_0)}{(F'(Q_0))^2} < 2$. Note also that $\lim_{Q_0 \to \infty} \Phi(Q_0) = \infty$ if $\lim_{Q_0 \to \infty} \frac{F(Q_0)}{F'(Q_0)} > -\infty$.

As can be seen from (9.4), clearly $u_k(\mathbf{a})$ is continuous in $\mathbf{a}$. Next, consider its first order derivative of $u_k(p_k) = \frac{W_k \log\left(1+\gamma_k^{MF}\right)}{1+e^{\lambda(I_0 - Q_0)}}$ w.r.t. $p_k$ where $\gamma_k^{MF}$ is given by (9.6): $\frac{\partial u_k(p_k)}{\partial p_k} = \frac{W e^{c_k p_k} g_k^{MF}(p_k)}{\left(1+\gamma_k^{MF}\right)\left(1+e^{\lambda\left(I_{0,-k}-Q_0\right)} e^{c_k p_k}\right)^2}$, where $c_k = \lambda \tilde{A}_k^2 \geq 0$ and $g_k^{MF}(p_k) = \frac{1}{N_k}\left(e^{-c_k p_k} + e^{\lambda\left(I_{0,-k}-Q_0\right)}\right)$
$- c_k e^{\lambda\left(I_{0,-k}-Q_0\right)}\left(1+\gamma_k^{MF}\right)\log\left(1+\gamma_k^{MF}\right)$. At an interior local extremum point for $p_k \in [0,\infty)$, we should have $g_k^{MF}(p_k) = \frac{1}{N_k} e^{-c_k p_k} + \frac{1}{N_k} e^{\lambda\left(I_{p,-k}-Q_0\right)}$
$- c_k e^{\lambda\left(I_{0,-k}-Q_0\right)}\left(1+\gamma_k^{MF}\right)\log\left(1+\gamma_k^{MF}\right) = 0$. Since $g_k^{MF}(0) = \frac{1}{N_k}\left(1 + e^{\lambda\left(I_{0,-k}-Q_0\right)}\right) > 0$, $g_k^{MF}(\infty) \longrightarrow -\infty$ and $g_k^{MF}(p_k)$ is continuous in $p_k$, clearly $g_k^{MF}(.)$ must have at least one zero crossing. However, since $g_k^{'MF}(p_k) = -\frac{c_k}{N_k}\left[e^{-c_k p_k} + e^{\lambda\left(I_{0,-k}-Q_0\right)}\left(1 + \log\left(1+\frac{p_k}{N_k}\right)\right)\right] < 0$ for $p_k \geq 0$, there is exactly one zero of $g_k^{MF}(p_k)$ on $[0,\infty)$, implying that $u_k(p_k)$ only has one local extremum point on $p_k \in [0,\infty)$. It follows that the local extremum point $p_k = p^*$

is indeed a global maximum of $u_k(.)$ on $[0,\infty)$, implying that $u_k(p_k)$ is quasi-concave in $p_k$, for each $k \in \mathcal{K}_s$.

From the above discussion it follows that the above game $G$ then has at least one Nash equilibrium due to the well-known result that NE exists in game $G = (\mathcal{K}, \mathcal{A}_k, u_k)$, if for all $k = 0, 1, \cdots, K_s$ (i) $\mathcal{A}_k$ is a non-empty, convex and compact subset of some Euclidean space $\mathbb{R}^N$, (ii) $u_k(\mathbf{p})$ is continuous in action vector $\mathbf{p}$, and (iii) $u_0(Q_0, \mathbf{a}_{-0})$ and $u_k(p_k, \mathbf{a}_{-k})$ are quasi-concave in $Q_0$ and $p_k$ respectively [106]. ■

# Appendix 9B

## Existence and Uniqueness of a NE with The Secondary LMMSE Receiver

Since secondary receivers don't influence the behavior of the primary system utility function, the quasi-concavity of the primary system utility function with the LMMSE secondary receiver still holds. Thus for the existence of a NE, the only condition that we need to establish anew is the quasi-concavity of secondary-user utility as a function of its power $p_k$, when the receiver is based on an LMMSE detector. We consider the first order derivative of $u_k(p_k, \mathbf{a}_{-k}) = \frac{W\log\left(1+\gamma_k^{MMSE}\right)}{1+e^{\lambda(I_0-Q_0)}}$ w.r.t. $p_k$, where $\gamma_k^{MMSE}$ is given by (9.7): $\frac{\partial}{\partial p_k}\{u_k(p_k)\} = \frac{We^{c_k p_k} g_k^{MMSE}(p_k)}{\left(1+\gamma_k^{MMSE}\right)\left(1+e^{\lambda(I_{0,-k}(\mathbf{p}_{-k})-Q_0)}e^{c_k p_k}\right)^2}$, where $c_k = \lambda \tilde{A}_k^2 \geq 0$ and $g_k^{MMSE}(p_k) = \frac{h_{kk}^2\left(\mathbf{s}_k^{(s)}\right)^T \Sigma_{k,k}^{-1}\mathbf{s}_k^{(s)}}{e^{c_k p_k}} + h_{kk}^2\left(\mathbf{s}_k^{(s)}\right)^T \Sigma_{k,k}^{-1}\mathbf{s}_k^{(s)} e^{\lambda(I_{0,-k}(\mathbf{p}_{-k})-Q_0)} - c_k e^{\lambda(I_{0,-k}(\mathbf{p}_{-k})-Q_0)}\left(1+\gamma_k^{MMSE}\right)\log\left(1+\gamma_k^{MMSE}\right)$. At an interior local extremum point for $p_k \in [0,\infty)$, we should have $g_k^{MMSE}(p_k) = 0$. Since $g_k^{MMSE}(0) = h_{kk}^2\left(\mathbf{s}_k^{(s)}\right)^T \Sigma_{k,k}^{-1}\mathbf{s}_k^{(s)}\left(1+e^{\lambda(I_{0,-k}(\mathbf{p}_{-k})-Q_0)}\right) > 0$, $g_k^{MMSE}(\infty) \longrightarrow -\infty$, and $g_k^{MMSE}(p_k)$ is continuous in $p_k$, function $g_k^{MMSE}(p_k)$ must have at least one zero crossing. However, since $g_k^{'MMSE}(p_k) = -c_k h_{kk}^2\left(\mathbf{s}_k^{(s)}\right)^T \Sigma_{k,k}^{-1}\mathbf{s}_k^{(s)}\left[e^{-c_k p_k} + e^{\lambda(I_{0,-k}(\mathbf{p}_{-k})-Q_0)}\left(1+\gamma_{k,k}^{(s)}\right)\right]$ $< 0$ for $p_k \geq 0$, there is exactly one zero of $g_{j,k}(p_k)$ on $[0,\infty)$, implying that $u_k(p_k)$ only has

one local extremum point on $p_k \in [0, \infty)$. Similar to the argument given in Appendix 9.1.7, it follows that $u_k(p_k)$ is indeed quasi-concave in $p_k$, for each $k \in \mathcal{K}_s$. Hence the C-DSL game has at least one Nash equilibrium due to the well known result of Debreu, Glicksberg and Fan [106].

To establish the uniqueness of the NE of the proposed C-DSL game with secondary LMMSE receiver, we show that the best response correspondence $r_k(\mathbf{a}_{-k})$ is a *standard function* for $k = 0, 1, \cdots, K_s$ [102]. For simplicity of exposition, below we assume $\bar{Q}_0 \longrightarrow \infty$ and $\bar{p}_k \longrightarrow \infty$. We adopt the convention that all the vector inequalities are component-wise.

**Primary system best response**

For $F(Q_0) = \log(1 + Q_0)$, the unique interior maximum of $u_0$ is given by

$$Q_0^* + (1 + Q_0^*) \log (1 + Q_0^*) \quad = \quad \left(\bar{Q}_0 + I_0\right). \tag{9.10}$$

Since $u_0 (Q_0)$ is monotonic increasing for $Q_0 < Q_0^*$, if the maximum interference cap is such that $\bar{Q}_0 < Q_0^*$, the best response is given by $r_0 (\mathbf{a}_{-0}) = \min\{\bar{Q}_0, Q_0^* (I_0)\}$.

I. *Positivity*: From (9.10), for $\mathbf{a}_{-0} = \mathbf{0}$, $Q_{0min}^* > 0$. So $r_0(\mathbf{a}_{-0}) > 0$ for $\mathbf{a}_{-0} \geq \mathbf{0}$.

II. *Monotonicity*: Since the left and the right hand side of (9.10) are increasing functions of $Q_0$ and $\mathbf{a}_{-0}$, respectively, given $\mathbf{a}_{-0} \geq \mathbf{a'}_{-0}$, $r_0(\mathbf{a}_{-0}) \geq r_0(\mathbf{a'}_{-0})$ .

III. *Scalability*: From (9.10), $Q_0^*(I_0)$ is concave in $I_0$ since $\frac{d^2 Q_0^*}{dI_0^2} = \frac{-1}{(1+Q_0^*)\left(2+\log(1+Q_0^*)\right)^3} < 0$ for $Q_0^* \geq 0$. It can be easily seen that positivity and concavity of $Q_0^*(I_0)$ together implies scalability. So for $\mu > 1$, we have $\mu r_0 (\mathbf{a}_{-0}) > r_0 (\mu \mathbf{a}_{-0})$.

Therefore, by Definition 3, the best response correspondence of the primary system is a standard function.

**Secondary links' best response**

The best response correspondence of the $k$-th secondary link is the transmit power which provides it with the optimum SINR $\gamma^{MMSE^*}$ given by the solution $p_k$ to the equation $g_k^{MMSE}(p_k) = 0$. Thus the best response correspondence of the $k$-th secondary link is $r_k(\mathbf{a}_{-k}) = \min\left\{\frac{\gamma^{MMSE^*} I_k^{(k)}}{h_{kk}^2}, \bar{p}_k\right\}$, where $I_k^{(k)} = \left[\left(\mathbf{s}_k^{(s)}\right)^T \Sigma_{k,k}^{-1} \mathbf{s}_k^{(s)}\right]^{-1}$. Since $\frac{\partial \gamma_k^{MMSE}(p_k)}{\partial p_k} = \frac{\gamma_k^{MMSE}}{p_k}$, maximizing the utility function for each user is equivalent to finding optimum SINR $\gamma^{MMSE*}$. Note also that $\gamma^{MMSE^*}$ is independent of $k$ as long as all secondary users have the same reward function.

I. *Positivity*: Since $\gamma^{MMSE^*} > 0$ and $I_k^{(k)} > 0$, the best response correspondence of the $k$-th secondary link $r_k(\mathbf{a}_{-k}) > 0$ for all $k \in \mathcal{K}_s$.

II. *Monotonicity*: By following a proof similar to [98], we have that for $\mathbf{a}_{-k} \geq \mathbf{a'}_{-k}$, $I_k^{(k)}(\mathbf{a}_{-k}) > I_k^{(k)}(\mathbf{a'}_{-k})$. Thus $r_k(\mathbf{a}_{-k}) = \frac{\gamma^{MMSE^*} I_k^{(k)}(\mathbf{a}_{-k})}{h_{kk}^2} \geq \frac{\gamma^{MMSE^*} I_k^{(k)}(\mathbf{a'}_{-k})}{h_{kk}^2} = r_k(\mathbf{a'}_{-k})$, for all $k \in \mathcal{K}_s$.

III. *Scalability*: For $\mu > 1$, $\mu r_k(\mathbf{a}_{-k}) = \frac{\mu \gamma^{MMSE^*} I_k^{(k)}(\mathbf{a}_{-k})}{h_{kk}^2}$ and $r_k(\mu \mathbf{a}_{-k}) = \frac{\gamma^{MMSE^*} I_k^{(k)}(\mu \mathbf{a}_{-k})}{h_{kk}^2}$. Similar to the proof given in [1], we have that $\mu I_k^{(k)}(\mathbf{a}_{-k}) > I_k^{(k)}(\mu \mathbf{a}_{-k})$. Hence, $\mu r_k(\mathbf{a}_{-k}) > r_k(\mu \mathbf{a}_{-k})$, for all $k \in \mathcal{K}_s$.

So, the noncooperative C-DSL game with secondary LMMSE receiver has a unique NE. ■

# Research Publications

The author has, so far, produced the following publications during his stay at the Communication and Information Sciences Laboratory (CISL), Department of Electrical and Computer Engineering, University of New Mexico:

- **K. Hakim**, M. Bkassiny and S. K. Jayaweera, "Dirichlet process mixture model (DPMM) based diffusive source localization using wireless sensor network (WSN)," *IEEE Transactions on Signal Processing*, Jan. 2013. (In review).

- **K. Hakim** and S. K. Jayaweera,"Source localization and tracking in a dispersive medium using wireless sensor network," *EURASIP Journal on Advances in Signal Processing*, July 2012. (under 2nd round review).

- **K. Hakim**, S. K. Jayaweera, G. El-howayek and C. Mosquera, "Efficient Dynamic Spectrum Sharing in Cognitive Radio Networks: Centralized Dynamic Spectrum Leasing (C-DSL)," *IEEE Transactions on Wireless Communications*, vol.9, no.9, pp.2956-2967, September 2010.

- **K. Hakim**, S. K. Jayaweera, G. El-Howayek, "Ad-hoc wireless sensor network based underwater diffusive source localization," *Fourth International Conference on Intelligent and Advanced Systems (ICIAS 12)*, Kuala Lumpur, Malaysia, June 2012.

- S. K. Jayaweera and **K. Hakim**, "A cooperative game theoretic solution for lifetime maximization of WSNs in sequential estimation," in *Proc. 5th International Conference*

*on Information and Automation for Sustainability (ICIAfS 2010)*, pp. 233-238, Dec. 2010.

- S. K. Jayaweera and **K. Hakim**, "Fairness in sequential estimation: A cooperative game theoretic solution for WSNs," in *Proc. IEEE 10th International Conference on Signal Processing (ICSP 2010)*, pp. 2427-2430, Oct. 2010.

- **K. Hakim**, S. K. Jayaweera, and C. Mosquera, "Analysis of Linear Receivers in a DSL Game for Spectrum Sharing in Cognitive Radio Networks," *2010 IEEE International Conference on Communications (ICC)*, pp.1-5, 23-27 May 2010.

- G. El-howayek, S. K. Jayaweera, **K. Hakim**, G. Vazquez-Vilar and C. Mosquera, "Dynamic Spectrum Leasing (DSL) in Dynamic Channels," *2010 IEEE International Conference on Communications (ICC) Workshops*, pp.1-5, 23-27 May 2010.

- S. K. Jayaweera and **K. Hakim**, "A Game-Theoretic Framework for Dynamic Spectrum Leasing (DSL) in Cognitive Radios," *2009 IEEE GLOBECOM Workshops*, pp.1-6, Nov. 30 2009-Dec. 4 2009.

# References


[1] "Gulf Coast Oil Disaster," Available at http://www.cnn.com/SPECIALS/2010/gulf.coast.oil.spill/.

[2] "BP Oil Spill," Available at http://www.guardian.co.uk/environment/bp-oil-spill.

[3] Arye Nehorai, Boaz Porat, and Eytan Paldi, "Detection and localization of vapor-emitting sources," *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 243 – 253, Jan. 1995.

[4] S. Vijayakumaran, Y. Levinbook, and T. F. Wong, "Maximum likelihood localization of a diffusive point source using binary observations," *IEEE Transactions on Signal Processing*, vol. 55, no. 2, pp. 665–675, Feb. 2007.

[5] Tong Zhao and Arye Nehorai, "Detecting and estimating biochemical dispersion of a moving source in a semi-infinite medium," *IEEE Transactions on Signal Processing*, vol. 54, no. 6, pp. 2213–2225, June 2006.

[6] Hiroshi Ishida, Takamichi Nakamoto, and Toyosaka Moriizumi, "Remote sensing and localization of gas/odor source and distribution using mobile sensing system," in *International Conference on Solid State Sensors and Actuators*, Chicago, IL, Jun. 1997, pp. 559–562.

[7] J. Patrick Fitch, Ellen Raber, and Dennis R. Imbro, "Technology challenges in responding to biological or chemical attacks in the civilian sector," *Science*, vol. 302, no. 5649, pp. 1350–1354, Nov. 2003.

[8] H. T. Banks and C. Castillo-Chavez, *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, Society for Industrial Mathematics, 2003.

[9] Aleksander Jeremic and Arye Nehorai, "Design of chemical sensor arrays for monitoring disposal sites on the ocean floor," *IEEE Journal of Oceanic Engineering*, vol. 23, no. 4, pp. 334 – 343, Oct. 1998.

[10] R. C. Hughes, G. C. Osboum, J. W. Bartholomew, and J. L. Rodriguez, "The detection of mixtures of NOx's with hydrogen using catalytic metal films on the sandia robust sensor with pattern recognition," in *The 8th International Conference on Solid-State Sensors and Actuators, and Eurosensors IX. Transducers '95*, Jun. 1995, pp. 730–733.

[11] James Weimer, Bruno Sinopoli, and Bruce H. Krogh, "Multiple source detection and localization in advection-diffusion processes using wireless sensor networks," in *Proc. 30th IEEE Real-Time Systems Symposium (RTSS)*, Philadelphia, PA, Dec. 2009, pp. 333–342.

[12] R. Bianchini and R. Rajamony, "Power and energy management for server systems," *IEEE Computer*, vol. 37, no. 11, pp. 68–76, Nov. 2004.

[13] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, 2002.

[14] M. Tubaishat and S. Madria, "Sensor networks: An overview," *IEEE Potentials*, vol. 22, no. 2, pp. 20–23, April 2003.

[15] Hans-Rolf Trankler and Olfa Kanoun, "Recent advances in sensor technology," in *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference (IMTC)*, May 2001, vol. 1, pp. 309–316.

[16] Takoi K. Hamrita, Nivedita P. Kaluskar, and Kurt L. Wolfe, "Advances in smart sensor technology," in *40th IAS Annual Meeting Industry, Industry Applications Conference*, Oct. 2005, vol. 3, pp. 2059–2062.

[17] Chao Hu Zhiyong, Liu Yingzi Pan, and Zhenxing Zeng, "A novel fpga-based wireless vision sensor node," in *Proceedings of the IEEE International Conference on Automation and Logistics Shenyang*, Shenyang, China, Aug. 2009, pp. 841–846.

[18] Mitchell S. Lebold, Brian Murphy, David Boylan, and Karl Reichard, "Wireless technology study and the use of smart sensors for intelligent control and automation," in *IEEE Aerospace Conference*, Big Sky, MT, Mar. 2005, pp. 1–15.

[19] "Micro-chemical sensors for in-situ monitoring and characterization of volatile contaminants," available at `http://www.sandia.gov/sensor/MainPage.htm`.

[20] "Mems micro hot plates suit chemical sensor applications," available at `http://news.thomasnet.com/fullstory/MEMS-Micro-Hot-Plates-suit-chemical-sensor-applications-818227`.

[21] "Cheaper chemical sensor," `http://www.technologyreview.com/tomarket/22861/`.

[22] "Li-cor underwater PAR sensors," http://www.lakescientist.com/2010/product-spotlight-li-cor-underwater-par-sensors.

[23] "Technologies for leak detection," available at http://www.sewerin.co.za/products/ex-tec-od-4.php.

[24] "Ac plug-in flammable gas and carbon monoxide detector (hs80504) alarm," available at http://testproducts.com/plugin-flammable-carbon-monoxide-detector-hs80504-alarm-p-60.html.

[25] "Amprobe gsd600 gas leak detector for methane and propane," available at http://www.testequipmentdepot.com/amprobe/envirotesters/gsd600.htm.

[26] "Ge security, 60-652-95, crystal carbon monoxide gas sensor," available at http://www.surveillance-video.com/60-652-95.html.

[27] "Waspmote," available at http://www.libelium.com/products/waspmote.

[28] J.C. Chen, Kung Yao, and R.E. Hudson, "Source localization and beamforming," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 30–39, Mar. 2002.

[29] P. Stoica and Jian Li, "Lecture notes - source localization from range-difference measurements," *IEEE Signal Processing Magazine*, vol. 23, no. 6, pp. 63–66, Nov. 2006.

[30] Mathias Ortner, Arye Nehorai, and Aleksander Jeremic, "Biochemical transport modeling and bayesian source estimation in realistic environments," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2520–2532, June 2007.

[31] Tong Zhao and Arye Nehorai, "Distributed sequential bayesian estimation of a diffusive source in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 55, no. 4, pp. 1511–1524, April 2007.

[32] Haotian Zhang, Jos M. F. Moura, and Bruce Krogh, "Dynamic field estimation using wireless sensor networks: Tradeoffs between estimation error and communication cost," *IEEE Transactions on Signal Processing*, vol. 57, no. 6, pp. 2383–2395, June 2009.

[33] Y.M. Lu, P.L. Dragotti, and M. Vetterli, "Localization of diffusive sources using spatiotemporal measurements," in *49th Annual Allerton Conference on Communication, Control, and Computing*, Sept. 2011, pp. 1072–1076.

[34] Xue Wang and Sheng Wang, "Collaborative signal processing for target tracking in distributed wireless sensor networks," *Journal of Parallel and Distributed Computing*, vol. 67, pp. 501–515, May 2007.

[35] Hui Ma and Brian W.H. Ng, "Collaborative data and information processing for target tracking in wireless sensor networks," in *Proc. of IEEE International Conference on Industrial Informatics*, Singapore, Aug. 2006, pp. 647–652.

[36] Dan Li, Kerry Wong, Yu H.Hu Hu, and Akbar Sayeed, "Detection, classification and tracking of targets in distributed sensor networks," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 17–29, Mar. 2002.

[37] Feng Zhao, Jaewon Shin, and James Reich, "Information-driven dynamic sensor collaboration for target tracking," *IEEE Signal Processing Magazine*, vol. 19, no. 2, pp. 61–72, Mar. 2002.

[38] O. Ozdemir, R. Niu, and P.K. Varshney, "Tracking in wireless sensor networks using particle filtering: Physical layer considerations," *IEEE Transactions on Signal Processing*, vol. 57, no. 5, pp. 1987–1999, May 2009.

[39] S. S. Ram and V. V. Veeravalli, "Localization and intensity tracking of diffusing point sources using sensor networks," *IEEE Global Telecommunications Conference (GLOBECOM)*, pp. 3107–3111, Nov. 2007.

[40] H. Vincent Poor, *An Introduction to Signal Detection and Estimation*, Springer, 1994.

[41] P.M. Djuric, J.H. Kotecha, Jianqui Zhang, Yufei Huang, T. Ghirmai, M.F. Bugallo, and J. Miguez, "Particle Filtering," *IEEE Signal Processing Magazine*, vol. 20, pp. 19–38, Sept. 2003.

[42] G. Roussas, "Nonparametric estimation in markov processes," *Annals of the Institute of Statistical Mathematics*, vol. 21, pp. 73–87, 1967.

[43] J.B. Predd, S.B. Kulkarni, and H.V. Poor, "Distributed learning in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 23, no. 4, pp. 56–69, Jul. 2006.

[44] V.N. Vapnik, "An overview of statistical learning theory," *IEEE Transactions on Neural Networks*, vol. 10, no. 5, pp. 988–999, Sep. 1999.

[45] Woojin Kim, Jaemann Park, and H.J. Kim, "Support vector learning approaches for object localization in acoustic wireless sensor networks," in *5th IEEE International Conference Intelligent Systems (ICIS)*, July 2010, pp. 485–489.

[46] Vin-sen Feng and Shih Yu Chang, "Determination of wireless networks parameters through parallel hierarchical support vector machines," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 3, pp. 505–512, Mar. 2012.

[47] N. Vasiloglou, A.G. Gray, and D.V. Anderson, "Parameter estimation for manifold learning, through density estimation," in *16th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing*, Sep. 2006, pp. 211–216.

[48] H. Wechsler, Z. Duric, Fayin Li, and V. Cherkassky, "Motion estimation using statistical learning theory," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 4, pp. 466–478, Apr. 2004.

[49] Filip Jurccek, Blaise Thomson, and Steve Young, "Reinforcement learning for parameter estimation in statistical spoken dialogue systems," *Computer Speech and Language*, vol. 26, no. 3, pp. 168–192, Oct. 2012.

[50] Renita Machado and Sirin Tekinay, "A survey of game-theoretic approaches in wireless sensor networks," *Computer Networks: The International Journal of Computer and Telecommunications Networking*, vol. 52, no. 16, pp. 3047–3061, Nov. 2008.

[51] Feng Zhao, Jie Liu, Juan Liu, Leonidas Guibas, and James Reich, "Collaborative signal and information processing: An information directed approach," in *Proceedings of the IEEE*, Aug. 2003, vol. 91, pp. 1199–1209.

[52] T. Peng W. Wang, Y. Cui and W. Wang, "Noncooperative power control game with exponential pricing for cognitive radio network," in *Vehicular Technology Conference (VTC2007-Spring)*, April 2007.

[53] Enrique Campos-Naòez, Alfredo Garcia, and Chenyang Li, "A game-theoretic approach to efficient power management in sensor networks," *Oper. Res.*, vol. 56, no. 3, pp. 552–561, May 2008.

[54] Shamik Sengupta and Mainak Chatterjee, "Distributed power control in sensor networks: A game theoretic approach," in *IWDC*, Dec. 2004, pp. 508–519.

[55] X. Zhang, Y. Cai, and H. Zhang, "A game theoretic dynamic power management policy on wireless sensor network," in *ICCT*, China, Nov. 2006, pp. 1–4.

[56] Liqiang Zhao, Hailin Zhang, and Jie Zhang, "Using incompletely cooperative game theory in wireless sensor networks," in *Proc. of IEEE WCNC*, Las Vegas, USA, Apr. 2008, pp. 1483–1488.

[57] Zhu Han, Zhu Ji, and K. J. Ray Liu, "Fair multiuser channel allocation for ofdma networks using nash bargaining solutions and coalitions," *IEEE Transactions on Communications*, vol. 53, no. 8, pp. 1366–1376, 2005.

[58] Ravi R. Mazumdar, Catherine Rosenberg, Senior Member, and Senior Member, "A game theoretic framework for bandwidth allocation and pricing in broadband networks," *IEEE/ACM Transactions on Networking*, vol. 8, pp. 667–678, 2000.

[59] Y. W. Teh, M. I. Jordan, M. J. Beal, and D. M. Blei, "Hierarchical Dirichlet processes," *Journal of American Statistical Association*, vol. 101, no. 476, pp. 1556–1581, 2006.

[60] Zhu Han and K. J. Ray Liu, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*, Cambridge University Press, 2008.

[61] G. Owen, *Game Theory*, Academic Press, Burlington, MA, 2001.

[62] Petr Tichavsky, Carlos H. Muravchik, and Arye Nehorai, "Posterior Cramer-Rao bounds for discrete-time nonlinear filtering," *IEEE Transactions on Signal Processing*, vol. 46, pp. 1386–1396, May 1998.

[63] R. Brown, "A brief account of microscopical observations made on the particles contained in the pollen of plants," *Edinburgh new Philosophical Journal*, pp. 358–371, 1828.

[64] J.N. Sherwood, *Diffusion processes*, In Proceedings of the Thomas Graham Memorial Symposium, University of Strathclyde. Gordon and Breach, 1971.

[65] J. Crank, *The Mathematics of Diffusion*, Oxford University Press, 1975.

[66] W. Jost, *Diffusion in solids, liquids, gases*, Academic Press Inc., 1952.

[67] Walter Rudin, *Principles of Mathematical Analysis*, McGraw-Hill, Inc., 1976.

[68] J. C. Lagarias, J. A. Reeds, M. H. Wright, and P. E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions," *SIAM Journal of Optimization*, vol. 9, pp. 112–147, 1998.

[69] Bruce Hoadley, "Asymptotic properties of maximum likelihood estimators for the independent not identically distributed case," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1977–1991, 1971.

[70] X. Liang, I. Balasingham, and S.-S. Byun, "A multi-agent reinforcement learning based routing protocol for wireless sensor networks," in *IEEE International Symposium on Wireless Communication Systems (ISWCS '08)*, Reykjavik, Iceland, Oct. 2008, pp. 552 –557.

[71] S.N. Simic, "A learning-theory approach to sensor networks," *IEEE Pervasive Computing*, vol. 2, no. 4, pp. 44 – 49, Oct.-Dec. 2003.

[72] P. Radivojac, U. Korad, K.M. Sivalingam, and Z. Obradovic, "Learning from class-imbalanced data in wireless sensor networks," in *IEEE 58th Vehicular Technology Conference (VTC 2003-Fall)*, Orlando, FL, Oct. 2003, vol. 5, pp. 3030 – 3034.

[73] S. Krishnamurthy, G. Thamilarasu, and C. Bauckhage, "Malady: A machine learning-based autonomous decision-making system for sensor networks," in *International Conference on Computational Science and Engineering (CSE '09)*, Vancouver, BC, Aug. 2009, vol. 2, pp. 93 –100.

[74] Y.W. Teh and M.I. Jordan, "Hierarchical bayesian nonparametric models with applications," *Bayesian nonparametrics*, pp. 158–207, 2010.

[75] Sun Li, Zhang Yanning, Ma Miao, and Tian Guangjian, "Sar image segmentation method using dp mixture models," in *International Symposium on Computer Science and Computational Technology (ISCSCT)*, Dec. 2008, vol. 2, pp. 598–601.

[76] Y.W. Teh, "A hierarchical bayesian language model based on pitman-yor processes," in *21st International Conference on Computational Linguistics*. Association for Computational Linguistics, 2006, pp. 985–992.

[77] M.D. Escobar and M. West, "Estimating normal means with a dirichlet process prior," *Journal of American Statistical Association*, vol. 89, no. 425, pp. 268–177, Mar. 1994.

[78] David Blackwell and James B. MacQueen, "Ferguson distributions via polya urn schemes," *The Annals of Statistics*, vol. 1, no. 2, pp. 353–355, Mar. 1973.

[79] M.D. Escobar and M. West, "Bayesian density estimation and inference using mixtures," *Journal of American Statistical Association*, vol. 90, no. 430, pp. 577–588, 1995.

[80] W. I. Tam and D. Hatzinakos, "An adaptive gaussian sum algorithm for radar tracking," in *IEEE Intl. Conf. on Communications 'Towards the Knowledge Millenium' (ICC)*, June 1997, vol. 3, pp. 1351–1355.

[81] N. J. Gordon, D. J. Salmond, and A. M. Smith, "Novel approach to non-linear/non-gaussian bayesian state estimation," in *Radar and Signal Processing, IEEE Proc. F*, June 1993, vol. 140, pp. 107–113.

[82] A. S. Chhetri, D. Morrell, and A. Papandreou-Suppappola, "Scheduling multiple sensors using particle filters in target tracking," in *IEEE Workshop on Statistical Signal Processing*, Sept. 2003, pp. 549–552.

[83] Sudharman K. Jayaweera and Carlos Mosquera, "Distributed sequential estimation with noisy, correlated observations," *IEEE Signal Processing Letters*, vol. 15, pp. 741–744, 2008.

[84] Thakshila Wimalajeewa and Sudharman K. Jayaweera, "Sequential estimation over noisy channels with distributed node selection," in *IEEE Military Communications Conference (MILCOM09)*, Boston, MA, USA, Oct. 2009.

[85] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*, Cambridge University Press, Mar. 2004.

[86] Marcel Lesieur, *Turbulence in fluids: Stochastic and numerical modeling*, Kluwer academic publishers, 1990.

[87] "Tokyo subways are attacked with sarin gas," available at `http://www.history.com/this-day-in-history/tokyo-subways-are-attacked-with-sarin-gas`.

[88] FCC, "Report of the spectrum efficiency working group," Tech. Rep., FCC spectrum policy task force, Nov. 2002.

[89] FCC, "ET docket no 03-322 notice of proposed rulemaking and order," Tech. Rep., Dec. 2003.

[90] D. I. Kim, L. B. Le, and E. Hossain, "Joint rate and power allocation for cognitive radios in dynamic spectrum access environment," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5517–5527, Dec. 2008.

[91] L. B. Le and E. Hossain, "Resource allocation for spectrum underlay in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5306–5315, Dec. 2008.

[92] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicenced bands," *IEEE Journ. Select Areas Commun.*, vol. 3, no. 25, pp. 517–528, Apr. 2007.

[93] Y. Xing, C. N. Mathur, M. A. Haleem, R. Chandramouli, and K. P. Subbalakshmi, "Dynamic spectrum access with QoS and interference temperature constraints," *IEEE Trans. Mobile Comp.*, vol. 6, no. 4, pp. 423–433, Apr. 2007.

[94] T. A. Weiss and F. K. Jondral, "Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency," *IEEE Commun. Mag.*, vol. 42, pp. 8–14, Mar. 2004.

[95] J. O'Daniell, *Analysis and Design of Cognitive Radio Networks and Distributed Radio Resource Management Algorithms*, Ph.D. thesis, Virginia Tech., 2006.

[96] S. K. Jayaweera and C. Mosquera, "A dynamic spectrum leasing (DSL) framework for spectrum sharing in cognitive radio networks," in *Proc. 43rd Annual Asilomar Conf. on Signals, Systems and Computers*, Pacific Grove, CA, Nov. 2009.

[97] J. Mitola, *Cognitive radio: An integrated agent architecture for software defined radio*, Ph.D. thesis, Royal Institute of Technology (KTH), Stockholm, Sweden, 2000.

[98] S. K. Jayaweera and T. Li, "Dynamic spectrum leasing in cognitive radio networks via primary-secondary user power control games," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 3300–3310, July 2009.

[99] S. K. Jayaweera, G. Vazquez-Vilar, and C. Mosquera, "Dynamic spectrum leasing: A new paradigm for spectrum sharing in cognitive radio networks," *IEEE Trans. Vehicular Technology*, vol. 59, no. 5, pp. 2328–2339, Jun. 2010.

[100] Sudharman K. Jayaweera, Kamrul Hakim, and Carlos Mosquera, "A game-theoretic framework for dynamic spectrum leasing (DSL) in cognitive radios," in *GLOBECOM Workshops, 2009 IEEE*, 30 2009-Dec. 4 2009, pp. 1–6.

[101] K. Hakim, S. K. Jayaweera, and C. Mosquera, "Analysis of linear receivers in a DSL game for spectrum sharing in cognitive radio networks," in *International Conference on Communications (ICC 2010)*, Cape Town, South Africa, May 2010.

[102] Roy D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1341–1347, 1996.

[103] G. El-Howayek and S. K. Jayaweera, "Distributed dynamic spectrum leasing (D-DSL) for spectrum sharing over multiple primary channels," *IEEE Trans. Wireless Commun.*, Feb. 2010, Submitted.

[104] S. Verdu, *Multiuser Detection*, Cambridge University Press, 1998.

[105] P. S. Maybeck, *Stochastic models, estimation, and control*, vol. 141 of *Mathematics in Science and Engineering*, U.K.: Academic, 1979.

[106] D. Fudenberg and J. Tirole, *Game Theory*, MIT Press, 1991.