

Multiple Transmitter Localization Using Received Signal Strength Measurements

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Jaime E. Almodovar
Master of Science
Johns Hopkins University, 2002
Bachelor of Science
University of Puerto Rico, Mayagüez, 1999

Director: Dr. Jill Nelson, Associate Professor
Department of Electrical and Computer Engineering

Fall Semester 2012
George Mason University
Fairfax, VA

Copyright © 2012 by Jaime E. Almodovar
All Rights Reserved

Dedication

I dedicate this dissertation to my family, especially my wife Carol and my daughters Isabella and Daniela.

Acknowledgments

I would like to thank all the people along the way that helped me get to this point; family members, mentors, and most importantly, God.

Table of Contents

	Page
List of Tables	viii
List of Figures	ix
Abstract	xii
1 Introduction	1
1.1 Background	1
1.2 RSS-based Localization	5
1.3 Related Work	7
1.4 Summary of Contributions	9
1.5 Thesis Outline	11
2 Interference Subtraction for Estimating the Location of Multiple Transmitters	12
2.1 System Model	12
2.2 Interference Analysis	14
2.2.1 Two Transmitter Analysis	15
2.2.2 General Case Analysis ($M > 2$)	21
2.3 Location Estimation of Multiple Transmitters	25
2.3.1 Maximum Likelihood Location Estimate of a Single Transmitter	26
2.3.2 Low-Complexity Interference Subtraction Algorithm	27
2.3.3 Interference Subtraction and the EM Algorithm	30
2.4 Transmitted Power Estimation	36
2.4.1 Incorporating Unknown Transmitted Power to the IS Estimation Framework	36
2.4.2 Estimating Multiple Transmitter Locations and Transmitted Powers via the EM algorithm	38
2.5 IS Performance Analysis	43
2.5.1 Expected Error with No Interference	43
2.5.2 Expected Error with Interference	45
2.5.3 Expected Error with Interference Subtraction	46
2.6 Simulation Results	50

2.6.1	Simulation Setup	50
2.6.2	Low-Complexity IS Algorithm	52
2.6.3	IS-based EM Algorithm	56
3	Distributed Processing for Multiple Transmitter Localization	60
3.1	Background and Related Work	60
3.2	Gossip Algorithm	63
3.3	Distributed MTL via the MTWARL Algorithm	64
3.4	Gossip Algorithm Convergence	67
3.4.1	Optimum Gossip Algorithm Configuration	68
3.4.2	Convergence Effects on Localization Performance	71
3.5	Energy Consumption Analysis	76
3.5.1	Energy Consumption of Centralized MTL Algorithms	77
3.5.2	Energy Consumption of the MTWARL Distributed MTL Algorithm	81
3.5.3	Energy Utilization Comparison Between Centralized and Distributed Approaches	84
3.6	Computational Complexity Analysis	87
3.6.1	Computational Complexity of Centralized MTL Algorithms	87
3.6.2	Computational Complexity of the MTWARL Distributed MTL Algorithm	89
3.6.3	Computational Complexity Comparison Between Centralized and Distributed Approaches	91
3.7	Localization Accuracy Comparison	91
4	Effects and Exploitation of Sensor Mobility in MTL	96
4.1	Extensions to System Model	96
4.1.1	Mobility Model	96
4.1.2	Correlated Shadowing Model	98
4.2	Implications of Correlated Shadowing to ML-based MTL Framework	99
4.3	Effects of Correlated Shadowing to Localization Performance	101
4.3.1	Insights from Prior Work on the Effects of Correlated Shadowing on RSS Localization	101
4.3.2	Effects of Shadowing Model Mismatch on the Localization Accuracy of the LCIS	103
4.4	Exploitation of Mobility in MTL	105
4.4.1	Extended LCIS for Mobility Exploitation	106
4.4.2	Simulation Results	110

4.5	Incorporating Refining Technique into the LCIS Framework	112
5	Conclusion and Future Work	118
5.1	Conclusion	118
5.2	Future Work	120
5.2.1	Algorithms for Estimating the Number of Transmitters	120
5.2.2	Alternative Gossip Algorithms for MTL	120
	Bibliography	122

List of Tables

Table		Page
2.1	Simulated performance of IS-EM algorithm for $M = 2$, $\sigma^2 = 36$ and $\gamma = 3$. .	48
3.1	Two transmitter energy consumption parameters per iteration.	83
3.2	Comparison of multiple transmitter localization algorithms.	95

List of Figures

Figure	Page
1.1 NTIA's frequency allocation chart for the United States.	3
2.1 Probability density function of $\frac{r_{2j}}{r_{1j}}$ and $1 + \frac{r_{2j}}{r_{1j}}$ for $\sigma = 6$, $\gamma = 3$ and $\frac{d_{1j}}{d_{2j}} = 2$	16
2.2 Probability density function of the interference noise ϵ	17
2.3 $\hat{\mathcal{E}}[\epsilon]$ as a function of $\frac{d_{1j}}{d_{2j}}$ for different Gauss-Laguerre integration orders, N ($\sigma_{int} = 2\sqrt{2}$ and $\gamma = 3$)	19
2.4 $\hat{V}[\epsilon]$ as a function of $\frac{d_{1j}}{d_{2j}}$ for different Gauss-Laguerre integration orders, N ($\sigma_{int} = 2\sqrt{2}$ and $\gamma = 3$)	20
2.5 Expected value and standard deviation of interference noise vs. distance ratio $\frac{d_{1j}}{d_{2j}}$ for $\gamma = 2$ (top), $\gamma = 3$ (middle), and $\gamma = 4$ (bottom).	22
2.6 Simulated and analytical approximation of $\mathcal{E}[\epsilon]$ for $M = 3$, $\sigma^2 = 36$, and $\gamma = 3$	25
2.7 Simulated PDF of the adjusted interference term $\epsilon_{mj} - \hat{\epsilon}_{mj}$ vs. Gaussian PDF with same mean and variance, and distance ratio equal to 0.6. $\hat{\epsilon}_{mj}$ was gen- erated given an independent, zero-mean, Gaussian distance ratio estimation error with standard deviation of 0.1.	34
2.8 Expected value and standard deviation of interference noise vs. power ratio $\frac{P_2^T}{P_1^T}$ for $\frac{d_{1j}}{d_{2j}} = 0.5$, $\gamma = 2$ (top), $\gamma = 3$ (middle), and $\gamma = 4$ (bottom).	39
2.9 Sensitivity of the expected noise with respect to the distance ratio $\frac{d_{1j}}{d_{2j}}$ for $\gamma = 2$ (top left), $\gamma = 3$ (top right), and $\gamma = 4$ (bottom).	51
2.10 Low-complexity IS algorithm two transmitter mean (left column) and median (right column) normalized squared distance error for $\sigma^2 = 16$ (top row), 36 (middle row) and 64 (bottom row).	53
2.11 Low-complexity IS algorithm three transmitter mean (left column) and me- dian (right column) normalized squared distance error for $\sigma^2 = 16$ (top row), 36 (middle row) and 64 (bottom row).	54

2.12	IS-based EM algorithm two transmitter mean (left column) and median (right column) normalized squared distance error for $\sigma^2 = 16$ (top row) and 36 (bottom row).	58
2.13	Extended IS-based EM algorithm mean (left column) and median (right column) normalized squared distance errors for $\sigma^2 = 16$ (top row) and 36 (bottom row).	59
3.1	Bounds on convergence time for $\psi = 0.05$ vs. number of receivers compared to the simulated number of iterations that yields the same ψ value.	71
3.2	Mean localization accuracy loss due to gossip convergence errors for different values of the gossip iteration factor F_g and $\sigma^2 = 36$	74
3.3	Two-transmitter mean localization performance of the MTWARL algorithm for different values of the gossip iteration factor F_g and $\sigma^2 = 36$	75
3.4	Cumulative distribution function of the convergence error E_t given a fixed number of algorithm iterations determined by different values of the gossip iteration factor F_g	76
3.5	Comparison of analytical and simulation PDF of $d(\phi_1, \phi_2)$ in a unit square area ($L = 1$).	80
3.6	Normalized average transmit energy of two versions of the MTWARL distributed localization algorithm compared to their centralized counterpart for $M = 2$ and $\gamma = 2$	85
3.7	Average transmit energy of the distributed algorithm version 2 relative to the energy consumption of the centralized algorithm for $\gamma = 2, 3, 4$, and 6.	86
3.8	Two-transmitter-localization normalized computational complexity comparison between centralized and distributed localization algorithms.	92
3.9	LCIS, IS-EM and MTWARL two transmitter mean (left column) and median (right column) localization accuracy for $\sigma^2 = 16$ (top row), 36 (middle row) and 64 (bottom row).	94
4.1	Two-transmitter mean performance comparison of LCIS algorithm (assuming uncorrelated shadowing) under uncorrelated and correlated shadowing operating in a dense urban propagation environment with parameters $\sigma^2 = 64$, $\gamma = 4$, $D_c = 20$ and $L_a = 1000$	105

4.2	Two-transmitter mean performance comparison of LCIS algorithm (assuming uncorrelated shadowing) under uncorrelated and correlated shadowing operating in a suburban propagation environment with parameters $\sigma^2 = 36$, $\gamma = 3$, $D_c = 500$ and $L_a = 10000$	106
4.3	Two-transmitter performance gain as a percentage of LCIS accuracy versus the number of receivers for various distance ratio thresholds in a propagation environment with $\sigma^2 = 36$ and $\gamma = 3$	111
4.4	Two-transmitter performance gain as a percentage of LCIS accuracy versus the number of receivers and percentage of total measurements used for estimate refinement for various distance ratio thresholds in a propagation environment with $\sigma^2 = 36$ and $\gamma = 3$	113
4.5	Two-transmitter performance gain as a percentage of LCIS accuracy versus the number of receivers and percentage of total measurements used for estimate refinement for various distance ratio thresholds in a propagation environment with $\sigma^2 = 16$ and $\gamma = 2$	114
4.6	Two-transmitter mean localization accuracy of the LCIS and modified LCIS versus the number of receivers under $\sigma^2 = 36$ and $\gamma = 3$	117

Abstract

MULTIPLE TRANSMITTER LOCALIZATION USING RECEIVED SIGNAL STRENGTH MEASUREMENTS

Jaime E. Almodovar, Ph.D.

George Mason University, 2012

Dissertation Director: Dr. Jill Nelson

Cognitive radio (CR) is often described as a context-aware intelligent radio capable of autonomous reconfiguration by learning from and adjusting to the radio environment. A critical source of information for a CR network is the locations of emitters operating in a particular frequency band, also known as primary transmitters. The use of this information has been shown to improve aspects of communication such as power efficiency, probability of interference, and data throughput. CR networks often estimate the location of primary transmitters based on uncoordinated measurements that contain co-channel interference. These measurements have different statistical properties than those without interference, and thus conventional localization techniques intended for clean measurements cannot be applied.

The work in this thesis explores the problem of using received signal strength (RSS) measurements taken by a network of CR nodes to estimate the locations of multiple primary transmitters that operate simultaneously in a given geographical area. A probabilistic model of the problem is developed, and algorithms to address location estimation challenges are proposed. Two approaches are proposed to solve the localization problem. The first approach is based upon approximating the maximum likelihood (ML) estimate of the

transmitter locations when no interference is present, and is computed via non-linear minimization at a centralized location. The second approach is based on the implementation of a proximity-based, closed-form, linear estimator computed in a distributed fashion across the network. The proposed algorithms provide a mechanism for CR users to obtain and/or update their information about nearby primary systems and offer a trade-off between computational complexity, fault tolerance, energy utilization and localization accuracy. Comparative analysis is performed in which all the trade off parameters are compared. This analysis enables the study of system level trades related to the implementation of multiple transmitter localization (MTL) algorithms. As a final extension of our work, the effects of sensor mobility on the statistical properties of RSS observations and its implications on ML-based MTL are studied. As part of this work, the proposed ML-based algorithm is extended to take advantage of additional measurements enabled by sensor mobility. In addition, we demonstrate that additional measurements can be used to iteratively reduce overall measurement uncertainty and thus improve localization performance.

Chapter 1: Introduction

1.1 Background

The emergence of new wireless technologies centered around wireless multimedia services has imposed new requirements on system capacity. These requirements come as a consequence of the need to deliver increased data rates to users. The advances in information theory of the last decades have made it possible to achieve channel information rates near Shannon's capacity limit by employing source and channel coding [1, 2]. This has allowed systems to increase capacity without increases in bandwidth. However, we have reached a point where the gap between Shannon's capacity and the capacity achieved by current coding techniques has significantly narrowed. Consequently, system designers have resorted to utilizing more bandwidth as the method for supporting the data rate increases required by the increasing demand. This need for more bandwidth has translated into growing demand for electromagnetic spectrum, a resource perceived as scarce given the current spectrum allocation scheme. Figure (1.1) shows the National Telecommunications and Information Administration (NTIA) frequency allocation chart. This chart shows a fixed overcrowded spectrum with little room for future growth. Faced with this problem, regulatory agencies have been forced to provide new ways to utilize the spectrum. In the efforts to further investigate the problem, several measurement campaigns have been undertaken. These measurements reveal that there are regions of the electromagnetic spectrum that are heavily underutilized [3]. Such findings put in question the suitability of the current spectrum allocation scheme for supporting current and future usage patterns. In addition, the studies discard the spectrum scarcity notion and suggest that it is more a spectrum management problem.

These developments have sparked major research activity from industry, academia, and

government. Efforts in the areas of engineering, economics and regulation have focused on ways of solving this spectrum efficiency problem. As studies have suggested, the problem lies in the fixed or static nature of the current spectrum allocation scheme. Therefore, in an attempt to solve the problem, the research community has focused on alternative schemes to the current static schemes. These alternatives are referred to as dynamic spectrum allocation schemes. As presented in [4], dynamic spectrum access strategies can be categorized under three broad models: 1) dynamic exclusive use, 2) open sharing, and 3) hierarchical access. These schemes vary in the degree to which they would alter the current allocation policy.

In the dynamic exclusive use model the current spectrum allocation policy is maintained, and added flexibility for the licensees is proposed in order to use the spectrum more efficiently. Under such schemes, the licensees would be given the right to sell and trade unused spectrum. An example of such a scheme would be a cell phone carrier selling or leasing the right to use some of their spectrum during night time when traffic is low and most of the spectrum lies idle. The main drawback of this type of scheme is that it does not assure a solution to the problem since sharing is not mandated by regulations [5].

In an open sharing model, open sharing is allowed among peer users as the basis for managing a spectral region. Under this model, different systems, each with its own set of objectives, co-exist in a spectral region as peers. This model draws support from the success of the industrial, scientific, and medical (ISM) frequency band in which wireless services can operate in an unlicensed fashion as long as they conform to the FCC operating rules for that band [6]. The challenge with this type of sharing model is designing operating rules that are efficient, fair and that do not incentivize selfish behavior [7]. As the success of the ISM band shows, schemes like these can work well (e.g. WiFi, ZigBee, and Bluetooth). However, since open sharing models are based on a flat hierarchy, their widespread implementation could encounter heavy opposition from licensed users with large operating infrastructures that would have to be modified as a consequence of such a policy move.

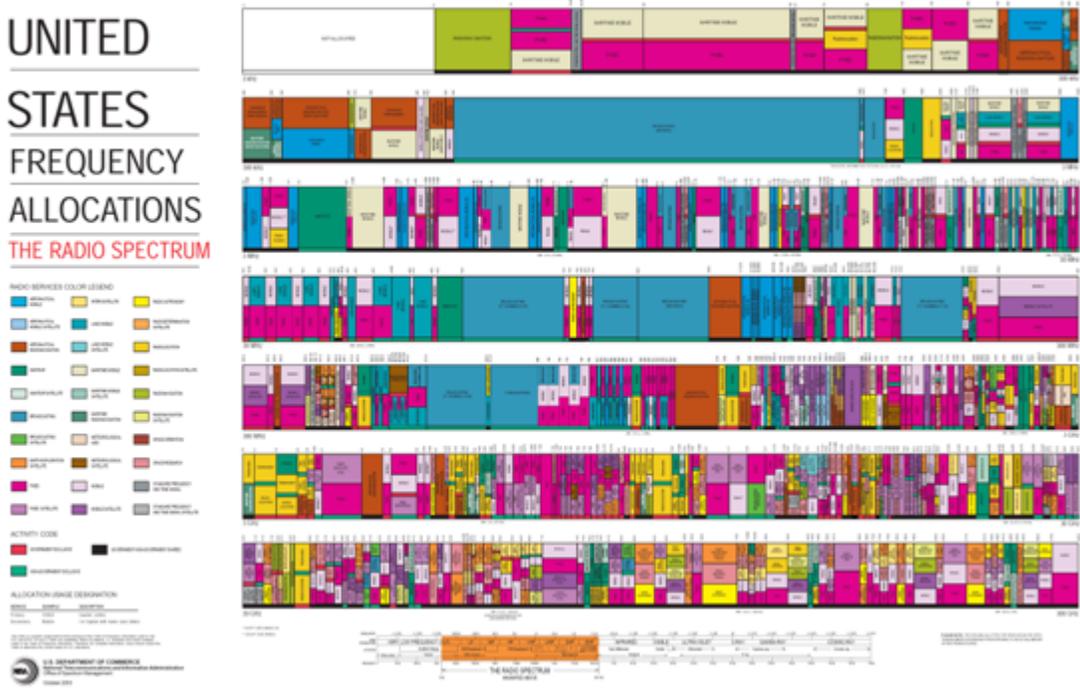


Figure 1.1: NTIA’s frequency allocation chart for the United States.

In the third category, the hierarchical access model, the concept of primary and secondary users is introduced. In this scheme, current licensees keep their rights to the spectrum and are considered the primary users. Primary users operate as usual without any modification to their infrastructure, while secondary users are permitted to access unused parts of the spectrum as long as they do not interfere with the primary users. Two main approaches have been considered for this scheme: spectrum underlay and spectrum overlay. In spectrum underlay, the secondary users utilize the spectrum by transmitting low spectral density waveforms over the transmissions of current primary users (e.g. direct sequence spread spectrum (DSSS) and ultra wideband (UWB) transmissions) [8,9]. This scheme imposes severe restrictions on the transmitted power of the secondary users because, in order to avoid interfering, secondary users’ transmissions must be below the noise floor of primary users’ receivers. In addition, this scheme does not directly target unused spectrum, as transmissions are spread equally over both used and unused frequencies. In contrast, spectrum overlay, also known as opportunistic spectrum access (OSA), attempts to target unused

spectrum directly in both time and space by having secondary users identify and exploit these spectral opportunities [10]. In this case, transmission constraints are determined on a case by case basis by the secondary users themselves, rather than by externally-imposed fixed global constraints. Because of its seeming advantages in addressing the spectrum efficiency problem directly and with minimal impact to the current infrastructure, spectrum overlay has received a great deal of attention from the research community, and significant advances have been made.

Alternative spectrum access and allocation strategies are frequently considered within the realm of Cognitive Radio. Cognitive radio (CR) was first proposed by Mitola [10] and was built upon the software-defined radio (SDR) paradigm. In his work, Mitola describes CR as a context-aware intelligent radio capable of autonomous reconfiguration by learning from and adjusting to the radio environment. The term cognitive radio is often used inaccurately when referring to dynamic spectrum access systems. However, the CR concept spans more than just spectral efficiency as it covers a much broader paradigm in which many aspects of a communication system can be improved by the use of cognition [11]. Power efficiency, probability of interception, and interference mitigation are examples of potential system characteristics that could be improved by the use of cognition. For example, if a wireless communication system knows the geographical locations and frequencies of its peers and of potential interferers, it could adapt using this information to avoid interfering or being interfered with by, for example, transmitting on a different frequency and/or using a directional antenna to transmit in the direction of its peer and away from the interfering radio.

In the work presented in this thesis, we consider CR in the context of spectral efficiency, more specifically CR networks that employ the spectrum overlay scheme of the hierarchical access model of dynamic spectrum access strategies. These types of systems are referred to as opportunistic spectrum access cognitive radio networks (OSACRNs). In OSACRNs, secondary users, also called CR nodes, sense the spectrum across relevant frequencies to determine when and with how much power they can transmit without causing harmful

interference to any primary (licensed) communication systems operating in the same area. This sensing process is referred to as spectral hole identification. Simple energy detection approaches have been proposed for spectrum hole identification. In such approaches, each node senses the electromagnetic energy in a particular frequency band and transmits only if the sensed power falls below a certain threshold [12]. Recent work has argued that these approaches fail to provide enough protection to the primary users against intolerable levels of interference [13]. It has also been argued that spectrum sensing can be significantly improved by incorporating knowledge about the primary system into the sensing process [14, 15]. Moreover, recent field tests have revealed that knowledge of the primary users' location and interference tolerance can dramatically improve spectrum sensing performance, and therefore reduce interference to the primary system [16].

One proposed method for smarter spectrum sensing is to maintain a database with information about the primary system; secondary users can then access the database to assist them in making their channel usage decisions [17]. This database can be populated by accessing FCC records or by storing estimates produced by the secondary network. We consider the problem of using received signal strength measurements taken by the secondary users to estimate the locations of multiple primary transmitters that operate simultaneously in adjacent coverage areas. We develop a probabilistic model of the problem and propose algorithms to address location estimation challenges. The proposed algorithms provide a mechanism for secondary users to obtain and/or update their information about nearby primary systems by processing distributed measurements. The efficiency of the algorithms is also of interest, and hence we explore complexity-performance tradeoffs as well as the distribution of computations over the network.

1.2 RSS-based Localization

Localization of an emitter based on measurements taken by a network of sensors is a problem that has been studied extensively in the area of wireless sensor networks (WSNs) [18–20].

Patwari et al. present a survey of work in this area [21]. In WSNs, measurements are taken in a collaborative (or coordinated) fashion in which the emitter to be located either transmits while the rest of the network listens, or it listens while a set of nodes transmits. On the contrary, in OSACRNs, the secondary users must take measurements in an uncoordinated fashion since the primary system and the secondary systems operate independently and do not exchange information. As a consequence of this lack of coordination, measurements taken by the secondary users often include energy from multiple emitters. This undesired energy distorts the measurements, which are expected to contain energy from a single source, and therefore reduce the accuracy of localization algorithms [22].

Patwari [21] describes the three main measurement types that have been considered for transmitter localization: time-of-arrival (TOA), angle-of-arrival (AOA) and received-signal-strength (RSS). We focus our work on RSS-based localization because of its relative simplicity compared to the alternative methods. RSS measurement systems do not require the precise synchronization or additional radio frequency (RF) hardware that TOA and AOA systems respectively require. In addition, computing average power is a byproduct of most spectrum sensing techniques, making the measurement readily available for other algorithms to use. These are highly desirable characteristics to CR systems as CR nodes are likely to be battery operated devices with limited power and computational resources.

When RSS measurements are used for multiple transmitter localization (MTL), each observation is comprised of the sum of the received powers from all of the active transmitters in the frequency band of interest. Each of these transmitted signals experiences random variations due to blockages, reflections, and scattering. Two types of models are often used for this type of impairment: 1) empirical, and 2) statistical. Empirical models are based on measurements taken over a particular range of distances in a given frequency range for a specific propagation environment [23, 24]. These measurements are then fitted into formulas to capture average path loss as a function of distance and frequency. Statistical models model the unknown characteristics of the propagation environment by defining path loss as a random variable [25]. This form of modeling allows for more fidelity, as it allows

the use of higher order statistics (e.g. variance) to define the propagation environment. In addition, statistical models enable the use of probabilistic estimation techniques, such as maximum likelihood, for parameter estimation.

In our work we assume a statistical model called lognormal shadowing to model the random path loss variations experienced by each transmitted signal [25, 26]. Lognormal shadowing is commonly used for this purpose and has been empirically confirmed in both indoor and outdoor propagation environments [27, 28]. Under this model, received power is assumed to be a random variable with a lognormal distribution. The lognormal model also assumes that the received signal is time-averaged for a long enough period, so that the small scale variations due to multi-path fading are removed.

MTL is a special case of the multiple source parameter estimation problem. This problem has been studied in the fields of underwater acoustics and array processing, but most of the work has focused on Gaussian noise models [14, 29–32]. In the case of RSS-based MTL, the observations are comprised of a sum of lognormal random variables. This sum of lognormal random variables does not yield a Gaussian distribution; hence conventional probabilistic approaches to the problem cannot be applied. Moreover, no closed-form probability density function exists to model this phenomenon, adding to the complexity of the problem.

1.3 Related Work

A modest amount of work has been done in the area of RSS-based localization in OS-ACRNs. Most of the work in this area has concentrated on localization when only a single transmitter is present in the sensor measurements. Mark and Nasif have studied the single transmitter localization problem in the context of spectral sensing and estimated the maximum interference-free transmit power (MIFTP) [15]. In their work, they demonstrate how knowledge of the primary transmitter’s location can be utilized to maximize transmission power while remaining below an allowed interference level. Dogandzic and Amran [33] have derived an EM solution to the single transmitter localization problem under a composite

gamma-lognormal channel model [34], but even in that case the solution is highly complex, requiring multivariate numerical integration. The use of the composite gamma-lognormal channel model in [33] attempts to combine the effects of both fading and shadowing on the observations of received power. Kim et al. employ a linear approximation to the relation between RSS measurements and transmitter location to frame a constrained optimization problem to estimate the transmitted power and location of a single transmitter under lognormal shadowing [35]. Their proposed approach involves a weighted least squares technique in which a weighting vector is derived to account for differences in shadowing seen by different sensors. Although Kim's approach offers a simple closed-form solution to the problem, it relies on a measurement model in which the shadowing is time-averaged at every sensor, rendering its use impractical in real world environments where shadowing noise is random in the spatial domain and cannot be averaged out over time. Nasif and Mark extended their work presented in [15] to develop an estimate of the MIFTP when multiple primary transmitters are present [22]. As part of this work, Mark and Nasif develop an approximated maximum likelihood estimate of transmitter locations under lognormal shadowing. Their approach estimates the location of a transmitter by maximizing an approximated likelihood function given knowledge about the interferer locations. Although the technique is shown to yield accurate results, the algorithm relies on obtaining estimates of the interferer locations from an external entity. In [36], Nelson et al. studied the MTL problem and derived a quasi EM approach to estimate locations under lognormal shadowing. The quasi EM technique is an iterative algorithm that estimates the location of multiple transmitters by, at each iteration, assigning a percentage of the received power to each transmitter and computing their location estimates independently. In their work, Nelson et al. show that given an adequate number of algorithm iterations and starting points, the algorithm can achieve a 3 dB accuracy gain over particle swarm optimization [37], a commonly used global optimization technique. In addition, the quasi EM algorithm requires no prior estimates of the interferer locations.

1.4 Summary of Contributions

As one of our contributions to the area of RSS-based localization in OSACRNs, we propose an estimation approach called Interference Subtraction (IS). The IS approach segments the MTL problem into independent single transmitter localization tasks. The algorithm approximates the maximum likelihood (ML) location estimate of a single transmitter by estimating the expected interference and removing it from the observations at each receiver. One of the primary objectives of the algorithm is to reduce computational complexity beyond that of existing algorithms [14, 33, 38], thus we employ the IS methodology to develop a low-complexity IS algorithm. The proposed algorithm outperforms competing algorithms by as much as 3 dB when computational complexity is constrained, thus offering the cognitive radio system designer a tradeoff between performance and computational complexity. The concept of IS can be applied to other estimation techniques. To demonstrate that, we propose the application of the IS method to develop an IS-based EM algorithm that iteratively estimates transmitter locations by estimating and subtracting the expected interference level at each receiver. The IS-based EM differs from the quasi EM of [36] in that the expectation step is performed based on subtraction of the interference from the observations, while the quasi EM performs it based on apportioning the total received power to each transmitter. The IS-based EM algorithm allows us to derive an approximation to the EM algorithm for solving the MTL problem, to which the conventional EM algorithm cannot be applied.

As part the IS framework, we have developed a novel interference analysis in which a closed-form approximation to the density function of the interference is derived. This model is crucial to the application of IS and could be useful in problems such as radio coverage, where co-channel interference is an issue.

To further our contribution, we have extended and applied prior work on information dissemination and linear single transmitter localization to develop a linear distributed MTL algorithm, namely the MTWARL algorithm. The algorithm utilizes a linear, closed-form, proximity-based location estimator, the weighted average receiver location, to iteratively

cluster and estimate transmitter locations. The algorithm employs a randomized pair-wise gossip algorithm to perform computations in a distributed fashion. The MTWARL algorithm provides an alternative localization algorithm that avoids the computational challenges inherent to ML approaches in which nonlinear minimizations are required to be solved. In addition, it offers the capability of performing computations distributedly, a feature that adds fault tolerance by eliminating single points of failure. We have shown that for large number of sensor nodes, the MTWARL algorithm can achieve greater localization accuracy than that of the IS-based centralized algorithms, while using as little as half of the energy in heavily obstructed communication environments. As part of this work, the convergence properties of the randomized pair-wise gossip algorithm are studied and the configuration that minimizes the lower bound on convergence time is derived. This result is critical to reducing the energy utilization of the MTWARL algorithm. Concluding our contribution to this topic, comparative analysis is performed in which the localization accuracy, energy utilization and computational complexity of the proposed distributed algorithm are compared to those of the competing centralized algorithms. This analysis enables the study of system level trades related to the implementation of MTL algorithms.

As a final contribution to RSS-based localization in OSACRN's, we have studied the effects of sensor mobility on the statistical properties of RSS observations and the implications that such behavior has on ML-based MTL. Our studies of these topics have revealed that the correlation of lognormal shadowing is the principal effect of this behavior on our system model. We have also found that the effects of correlated shadowing on MTL are minimal and that the gains obtained by additional processing to account for it are too small to justify large increases in computational complexity. As part of this work, the LCIS algorithm has been extended to take advantage of the additional measurements enabled by sensor mobility. We have demonstrated that additional measurements can be used to iteratively reduce overall measurement uncertainty and thus improve localization performance.

1.5 Thesis Outline

The remainder of the thesis is organized as follows. In Chapter 2, we delve in detail into the MTL problem and present a framework that leads to a proposed approach to address the problem. We begin by defining a system model. An interference analysis follows in which a statistical model is developed for the measurement interference. We then present the proposed localization algorithms, both of which are based on a novel technique called interference subtraction (IS). Following that, we extend one of the proposed algorithms to jointly estimate location and transmitted power. We continue the chapter with a performance analysis of the IS technique in which we demonstrate the mechanisms by which it reduces the effects of the interference, thus improving location estimation accuracy. We conclude the chapter with simulation results of the algorithms presented. In Chapter 3, the work of Chapter 2 is extended and a linear distributed approach to MTL is explored. We begin the chapter with background information on the topic. In the subsequent section, the proposed distributed algorithm, the MTWARL, is presented in detail. Following that, we analyze the convergence of the gossip algorithm employed by the MTWARL algorithm and derive the gossip configuration that minimizes the lower bound on convergence time. We conclude the chapter with a comparative analysis in which the localization accuracy, energy utilization and computational complexity of the proposed distributed algorithm are compared to those of the competing centralized algorithms. In Chapter 4, we explore the effects of sensor mobility to MTL. We start the chapter presenting how the system model is extended to reflect this behavior. We continue the chapter with an analysis of the implications that those system model changes have on ML-based MTL. Following that, we present a treatment of the effects of correlated shadowing to localization accuracy. The chapter is concluded with a discussion on exploiting sensor mobility to increase localization accuracy. Chapter 5 finalizes the thesis with discussions on conclusions and future work.

Chapter 2: Interference Subtraction for Estimating the Locations of Multiple Transmitters

In this chapter, we delve in detail into the multiple transmitter localization problem and propose two algorithms to address it, both of which are based on a novel technique called interference subtraction (IS). We start the chapter by defining a system model and performing an analysis of the measurement interference, which is defined as the portion of the received power coming from transmitters other than the one of interest. We then present the proposed algorithms, which are followed by a performance analysis and simulation results.

2.1 System Model

Consider M primary transmitters and N cognitive radio nodes (also referred to as sensors or receivers) located within a square region of arbitrary area. We assume that the locations of the primary transmitters are unknown and are denoted by $\theta = [\theta_1 \ \theta_2 \ \dots \ \theta_M]^T$, where θ_i denotes the two-dimensional location of the i th transmitter. The locations of the N sensors are assumed to be known but arbitrary. The cognitive radio nodes may be affixed, for example, to vehicles or to individuals, and hence their locations will be defined by the activities of the “carriers” rather than by the best geometry for obtaining reliable transmitter location estimates.

We assume a log-distance path loss model [39] such that the noise-free received power at the j th receiver from the i th transmitter is given by

$$P_{ij} = P_i^T \rho \left(\frac{d_0}{d_{ij}} \right)^\gamma, \quad (2.1)$$

where P_i^T is the power transmitted by the i th transmitter, ρ is a constant set by the frequency of operation and antenna properties, d_0 is a reference close-in distance from the i th primary transmitter, d_{ij} is the two-dimensional Euclidean distance from the i th transmitter to the j th receiver, and γ is the path loss exponent. The log-distance model is a simplified linear (in the log domain) piecewise model, with two segments, where free-space propagation is assumed up to distance d_0 and non-free-space propagation is assumed for distances greater than d_0 . Typical values for the close-in distance are 1-10 meters for indoor environments and 10-100 meters for outdoor environments. The path loss exponent γ sets the rate at which the signal attenuates with distance; its value depends on the propagation environment. Typical values of γ range between 2, for free-space propagation, and 6, for heavily cluttered urban environments [25].

Let $r_{(dB)} = [r_{1(dB)} \ r_{2(dB)} \ \dots \ r_{N(dB)}]^T$, where $r_{j(dB)}$ denotes the observed power at the j th cognitive radio node measured in dB. Each received power observation $r_{j(dB)}$ represents an aggregate of the energy emanating from all primary transmitters in the cognitive radio node's vicinity. The contribution of each primary transmitter to $r_{j(dB)}$ is affected by obstructions and clutter according to a lognormal shadowing model. This relation can be expressed as

$$\begin{aligned} r_{j(dB)} &= 10 \log_{10} \left(\sum_{i=1}^M r_{ij} \right) \\ &= 10 \log_{10} \left(\sum_{i=1}^M P_{ij} \kappa_{ij} \right), \end{aligned} \tag{2.2}$$

where r_{ij} is power received at the j th node from the i th primary transmitter measured in linear units, and $\kappa_{ij} = 10^{\frac{X_{ij}}{10}}$ is the lognormal shadowing noise affecting each transmitter-node link. Each X_{ij} is assumed to have a Gaussian distribution with mean 0 and variance σ_{ij}^2 . We assume independent and homogeneous lognormal shadowing across all nodes, thus $\sigma_{ij}^2 = \sigma^2$. The independence assumption is valid as long as there is sufficient separation between nodes

to ensure that the signals observed at each node undergo independent obstruction effects [40]. In addition, all transmitters are assumed to transmit with equal known powers, thus $P_i^T = P^T$ is a known constant. This assumption is carried throughout the first sections of this chapter for the purpose of narrowing the scope of the work. In practice however, estimation of transmitted power would be required. This point is addressed in Section 2.4, where this assumption is removed and estimation of transmitted power is incorporated into the localization problem.

When considering the problem of estimating transmitter locations in an environment where multiple transmitters are concurrently active, it is helpful to express the measurements in terms of desired and interfering quantities. With that in mind, we can designate one of the transmitters as the desired transmitter and rearrange equation (2.2) as follows:

$$\begin{aligned} r_{j(dB)} &= 10 \log_{10}(r_{mj}) + 10 \log_{10} \left(1 + \sum_{i=1, i \neq d}^M \frac{r_{ij}}{r_{mj}} \right) \\ &= r_{mj(dB)} + \epsilon_{mj}, \end{aligned} \tag{2.3}$$

where $r_{mj(dB)} = 10 \log_{10}(r_{mj})$, $m \in [1, 2, \dots, M]$ is the received power at the j th node from the desired or main primary transmitter, and $\epsilon_{mj} = 10 \log_{10}(1 + \sum_{i=1, i \neq d}^M \frac{r_{ij}}{r_{mj}})$ is the undesired or interference term affecting the measurement of received power from m th transmitter. As equation (2.3) shows, the interference term ϵ_{mj} is additive, a desired property in the development of our localization techniques. This interference term will be referred to as the interference noise throughout this dissertation.

2.2 Interference Analysis

In this section, we perform statistical analysis of the interference noise with the objective of gaining insight into its behavior and its effect on the transmitter localization problem. For simplicity, we start our analysis with the two primary transmitter case ($M = 2$). The

analysis is then extended to a general case where more than two primary transmitters are present.

2.2.1 Two Transmitter Analysis

To understand the effects of the interference on the estimation process, we study the statistics of the interference noise ϵ_{mj} . We first consider its probability density function (pdf). As defined in (2.3), when two transmitters are present, the interference noise affecting the received power contribution from transmitter 1 ($m = 1$) measured at the j th receiver is given by

$$\begin{aligned}\epsilon_{1j} &= 10 \log_{10} \left(1 + \frac{r_{2j}}{r_{1j}} \right) \\ &= 10 \log_{10} \left(1 + \left(\frac{d_{1j}}{d_{2j}} \right)^\gamma 10^{\frac{(X_{2j} - X_{1j})}{10}} \right).\end{aligned}\quad (2.4)$$

The interference affecting the measurement from primary transmitter 2, ϵ_{2j} , can be computed in similar manner. Since ϵ_{1j} is the log of the random variable $1 + \frac{r_{2j}}{r_{1j}}$ and the distribution of $1 + \frac{r_{2j}}{r_{1j}}$ can be determined, the distribution of ϵ_{1j} can be straightforwardly obtained by applying geometric reconstruction [41]. To determine the distribution of the term $1 + \frac{r_{2j}}{r_{1j}}$, we note that $\frac{r_{2j}}{r_{1j}}$ is a lognormal random variable with parameters given by

$$\mu_{int} = 10\gamma \log_{10} \left(\frac{d_{1j}}{d_{2j}} \right) \quad \sigma_{int} = \sqrt{2}\sigma, \quad (2.5)$$

where μ_{int} and σ_{int} are the mean and variance, respectively, of the Gaussian distribution characterizing the term $10 \log_{10} \left(\frac{r_{2j}}{r_{1j}} \right)$. Thus, the distribution of $1 + \frac{r_{2j}}{r_{1j}}$ is a shifted version of the lognormal distribution that characterizes $\frac{r_{2j}}{r_{1j}}$, in which the mean is augmented by 1 as shown in Figure 2.1.

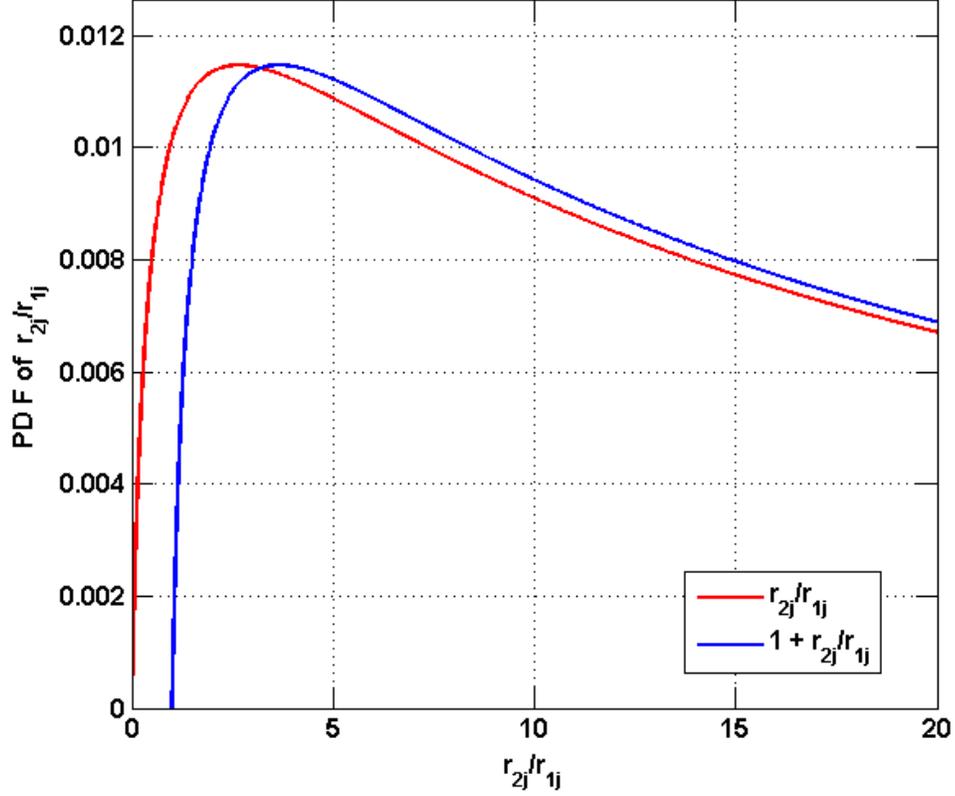


Figure 2.1: Probability density function of $\frac{r_{2j}}{r_{1j}}$ and $1 + \frac{r_{2j}}{r_{1j}}$ for $\sigma = 6$, $\gamma = 3$ and $\frac{d_{1j}}{d_{2j}} = 2$.

Given the above information, we can apply geometric reconstruction to obtain the pdf of ϵ_{1j} as

$$f_{\epsilon}(\epsilon) = \frac{10^{\epsilon/10}}{\sigma_{int} \sqrt{2\pi} (10^{\epsilon/10} - 1)} \exp \left(\frac{-(10 \log_{10}(10^{\epsilon/10} - 1) - 10\gamma \log_{10}(\frac{d_{1j}}{d_{2j}}))^2}{2\sigma_{int}^2} \right). \quad (2.6)$$

Note that for clarity we dropped the subscripts and use ϵ to denote the interference noise random variable. Equation (2.6) shows that with equal transmitted powers, the pdf of ϵ_{1j} has three parameters: $\frac{d_{1j}}{d_{2j}}$, σ_{int}^2 , and γ . It also shows that, unlike the statistics of shadowing noise in a homogeneous environment, the statistics of the interference noise at

each cognitive radio node vary depending on the CR node location relative to the desired and interfering transmitter locations. This information suggests that one could reduce the effects of interference by configuring sensors in a way that minimizes interference noise. Figure 2.2 shows a set of pdfs corresponding to the interference noise ϵ for different distance ratios $\frac{d_{1j}}{d_{2j}}$, $\sigma_{int}^2 = 16$ and $\gamma = 3$. The figure illustrates that in contrast to shadowing noise,

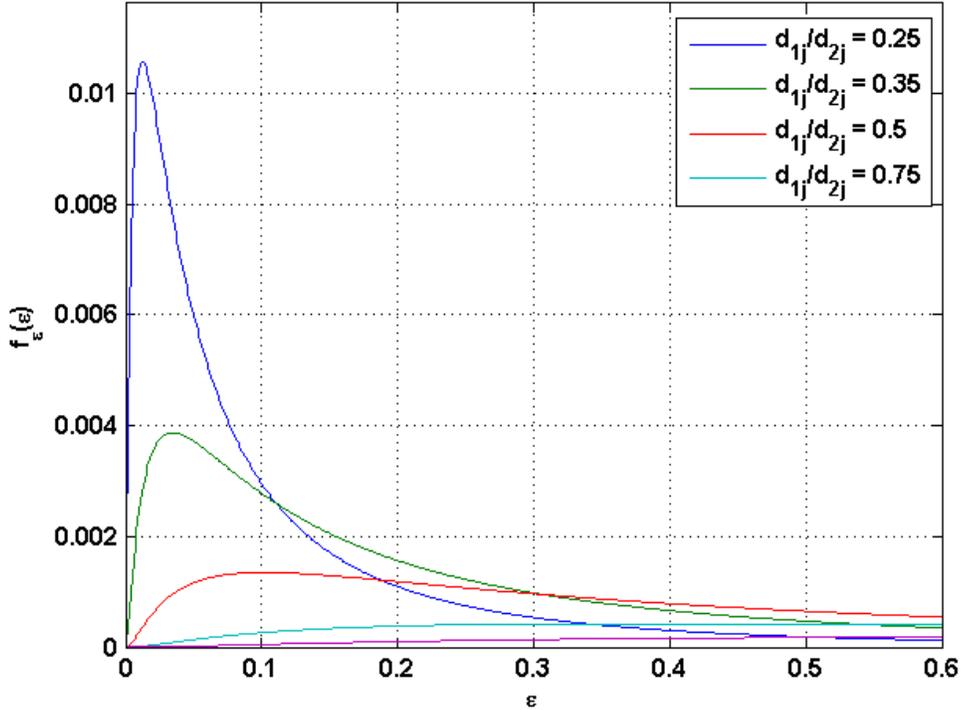


Figure 2.2: Probability density function of the interference noise ϵ

the interference noise term is always positive. This is expected, since the interference is undesired signal power from other receivers. Figure 2.2 also shows that the distribution of the interference noise is highly dependent on the distance ratio $\frac{d_{1j}}{d_{2j}}$. This quantifies what one would expect, that when the main transmitter is much closer to a receiver than the interferer (a smaller distance ratio), the interference noise density concentrates near zero and has a small spread. On the other hand, when the interferer is much closer to a receiver than the main transmitter (a larger distance ratio) the interference noise pdf flattens as its

variance increases.

Now that we have an expression for $f_\epsilon(\epsilon)$, we can derive expressions for the interference noise statistics. In our analysis we are interested in the expected value and variance, as they provide valuable insight into the behavior of the interference noise. Since the expression for the density function (2.6) is not integrable, we compute these statistics by numerical integration, using Gauss-Laguerre (GL) quadrature [42] to approximate the expectation integrals. The expression for the GL quadrature approximation of the mean of the interference noise is given by

$$\begin{aligned} \mathcal{E}[\epsilon] &= \int_0^\infty \epsilon f_\epsilon(\epsilon) d\epsilon \\ &= \int_0^\infty \exp(-\epsilon) [\exp(\epsilon)\epsilon f_\epsilon(\epsilon)] d\epsilon \\ &= \sum_{n=1}^N w_n \exp(a_n) a_n f_\epsilon(a_n) + R_N, \end{aligned} \quad (2.7)$$

where a_n are the abscissas or nodes, w_n are the weights or coefficients, and N is the integration order. GL quadrature is an extension of the Gaussian quadrature method for integrals of the form $\int_0^\infty \exp(-x)f(x)dx$ where the roots of Laguerre polynomials are used as abscissas [42]. The abscissas and weights for GL quadrature formulas can be found in [43]. Substituting equation (2.6) into (2.7), and removing the remainder R_N , we define an estimate of the mean of ϵ , $\hat{\mathcal{E}}[\epsilon]$, as follows

$$\hat{\mathcal{E}}[\epsilon] = \sum_{n=1}^N \frac{w_n \exp(a_n) a_n 10^{a_n/10}}{\sigma_{int} \sqrt{2\pi} (10^{a_n/10} - 1)} \exp\left(\frac{-(10 \log_{10}(10^{a_n/10} - 1) - 10\gamma \log_{10}(\frac{d_{1j}}{d_{2j}}))^2}{2\sigma_{int}^2}\right). \quad (2.8)$$

Figure 2.3 shows the impact of N on the accuracy of $\hat{\mathcal{E}}[\epsilon]$. This is accomplished by comparing $\hat{\mathcal{E}}[\epsilon]$ to the results obtained from a Monte Carlo simulation of the interference noise ϵ . As

the figure shows, $\mathcal{E}[\epsilon]$ can be approximated with negligible approximation error by $\hat{\mathcal{E}}[\epsilon]$ with $N = 20$.

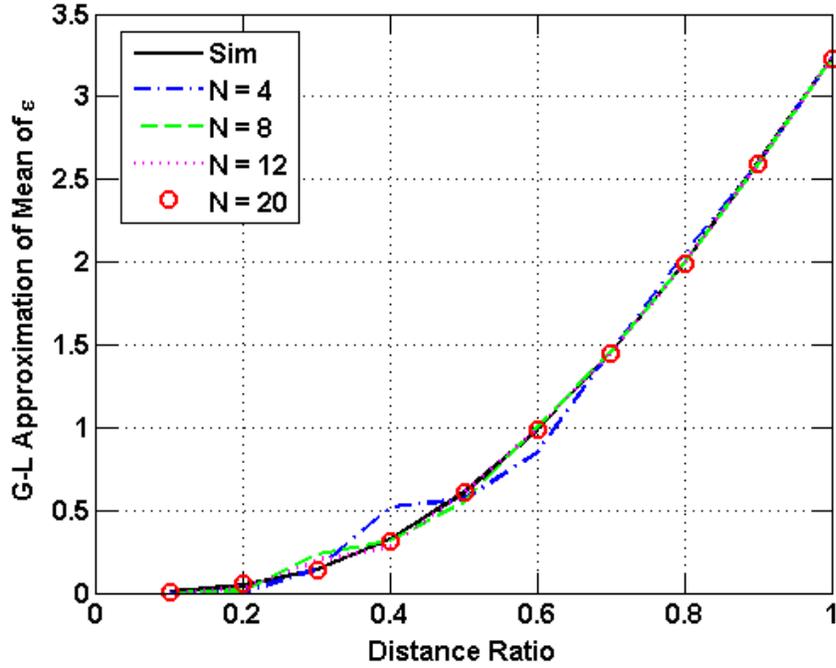


Figure 2.3: $\hat{\mathcal{E}}[\epsilon]$ as a function of $\frac{d_{1j}}{d_{2j}}$ for different Gauss-Laguerre integration orders, N ($\sigma_{int} = 2\sqrt{2}$ and $\gamma = 3$)

Similarly for the variance of ϵ , we can express its GL quadrature approximation and estimate as follows

$$\begin{aligned}
 V[\epsilon] &= \int_0^\infty (\epsilon - \mathcal{E}[\epsilon])^2 f_\epsilon(\epsilon) d\epsilon \\
 &= \sum_{n=1}^N w_n \exp(a_n) (a_n - \mathcal{E}[\epsilon])^2 f_\epsilon(a_n) + R_N,
 \end{aligned} \tag{2.9}$$

$$\hat{V}[\epsilon] = \sum_{n=1}^N \frac{w_n \exp(a_n) (a_n - \hat{\mathcal{E}}[\epsilon])^2 10^{a_n/10}}{\sigma_{int} \sqrt{2\pi} (10^{a_n/10} - 1)} \exp\left(\frac{-(10 \log_{10}(10^{a_n/10} - 1) - 10\gamma \log_{10}(\frac{d_{1j}}{d_{2j}}))^2}{2\sigma_{int}^2}\right). \tag{2.10}$$

Figure 2.4 shows the impact of N on the accuracy of $\hat{V}[\epsilon]$, which is also accomplished by comparing $\hat{V}[\epsilon]$ to the results obtained from a Monte Carlo simulation of the interference noise ϵ . As the figure shows, $V[\epsilon]$ can also be approximated with negligible approximation error by $\hat{V}[\epsilon]$ with $N = 20$.

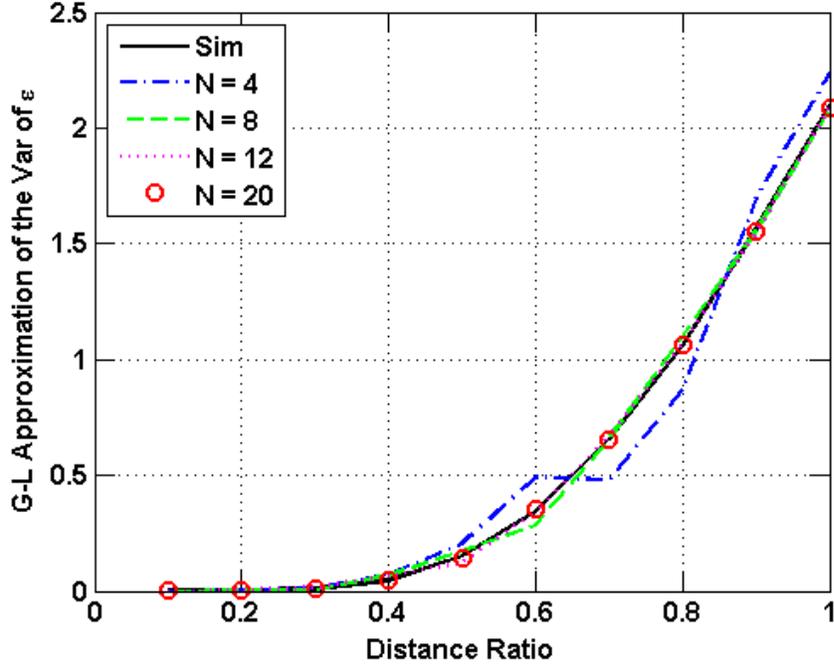


Figure 2.4: $\hat{V}[\epsilon]$ as a function of $\frac{d_{1j}}{d_{2j}}$ for different Gauss-Laguerre integration orders, N ($\sigma_{int} = 2\sqrt{2}$ and $\gamma = 3$)

Figure 2.5 shows the computed statistics of the interference noise term ϵ versus distance ratio for different values of σ_{int} and γ . Each row of plots corresponds to a value of γ , and the columns correspond to the expected value and the standard deviation of ϵ , respectively. The range of values plotted allows us to study the behavior of the interference in a variety of propagation environments. The low values of γ and σ_{int} correspond to rural areas with little propagation obstructions, whereas high values of these parameters represent urban areas with significant propagation obstructions [39]. Concurrent low values of γ and high values of σ_{int} and vice versa are unlikely in practice but are presented for completeness. The

plots also confirm what the pdf in Figure 2.2 suggests: as $\frac{d_{1j}}{d_{2j}}$ increases, both the expected value and standard deviation of the interference noise increase. Another relation evident in the plots is that the expected value and standard deviation of the interference noise increase with larger σ_{int} . This is also expected as the more uncertainty the channel introduces, the more uncertainty the interference will have. Finally, it is noted that the path loss exponent γ affects the values of the interference noise statistics. This is evident at low distance ratios where both the mean and standard deviation of the interference noise decrease with larger γ . However, this effect dissipates as the distance ratio reaches one. This gradual decrease in the effect of γ on the statistics of ϵ is due to the power decay of electromagnetic waves as they propagate through the wireless channel, which is proportional to $d^{-\gamma}$. For larger values of γ , the transmitted signals undergo much more attenuation before arriving at the receiver, and therefore the expected interference and interference shadowing effects are smaller. When $\frac{d_{1j}}{d_{2j}} = 1$, signal power from both the main and interfering transmitters experience the same path loss and therefore γ has no effect on the interference statistics.

2.2.2 General Case Analysis ($M > 2$)

The analysis for $M > 2$ proceeds similarly to that for $M = 2$. The main difference between the two is that the pdf of the interference noise $f_\epsilon(\epsilon)$ is estimated rather than derived in closed-form. Once $f_\epsilon(\epsilon)$ is estimated, the same numerical procedure is used to estimate the mean and variance of ϵ .

As defined in equation (2.3), the interference noise affecting the received power measurements from the m th transmitter taken by the j th receiver when M transmitters are present is given by

$$\epsilon_{mj} = 10 \log_{10} \left(1 + \sum_{i=1, i \neq m}^M \frac{r_{ij}}{r_{mj}} \right), \quad (2.11)$$

where m is the subscript that identifies the main or desired primary transmitter to the

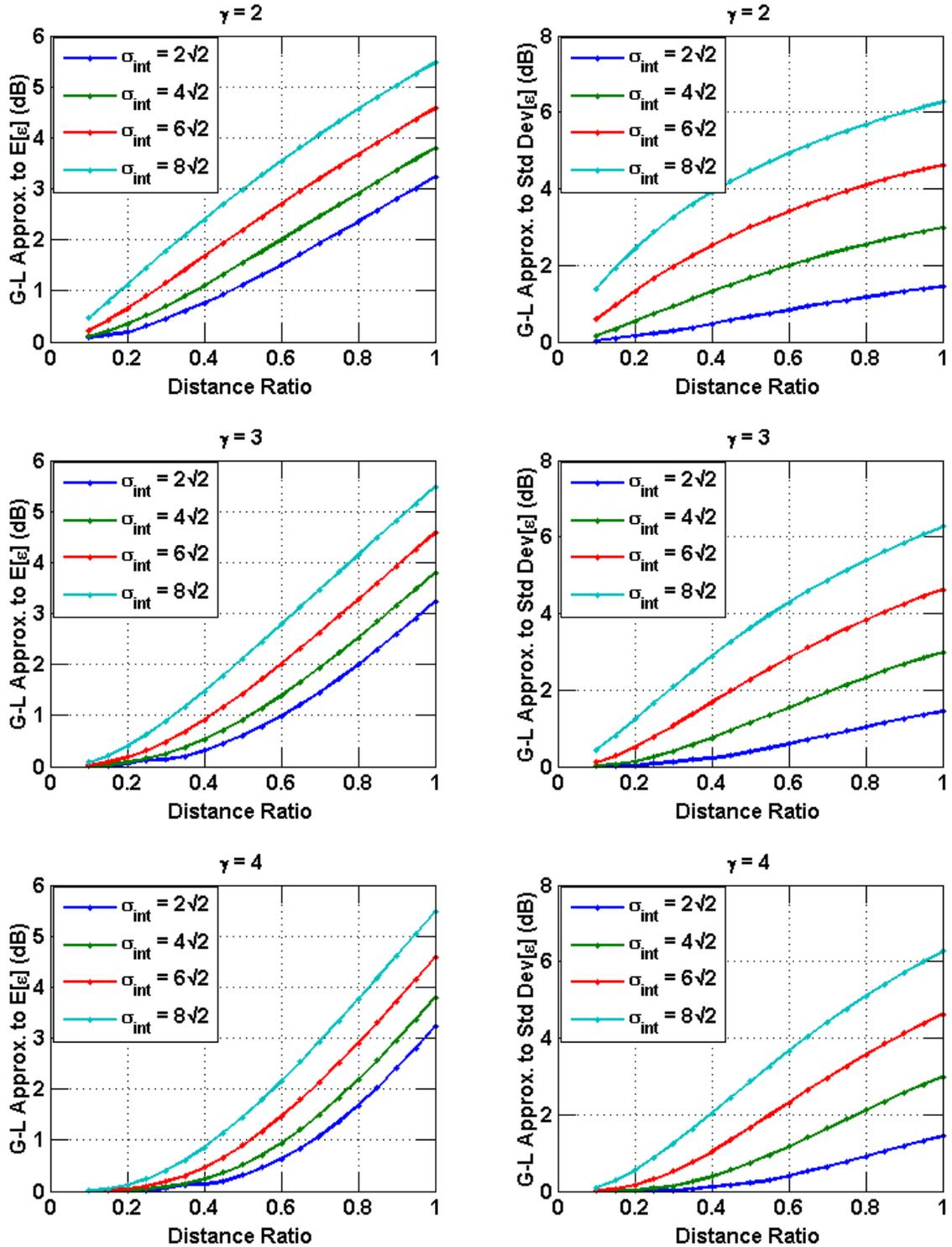


Figure 2.5: Expected value and standard deviation of interference noise vs. distance ratio $\frac{d_{1j}}{d_{2j}}$ for $\gamma = 2$ (top), $\gamma = 3$ (middle), and $\gamma = 4$ (bottom).

receivers estimating its location. Equation (2.11) shows that inside the logarithm is a summation of $(M - 1)$ lognormal random variables. Since no analytic distribution exists for the sum of lognormals, we use the Fenton-Wilkinson method [44] to approximate the sum of lognormal random variables by a single lognormal random variable z with different mean and variance that are determined by the parameters of the individual lognormals in the sum. Using this approach, the pdf of ϵ_j can be approximated in the form of equation (2.6), which allows the computation of the expectation integral to obtain $\hat{\mathcal{E}}[\epsilon]$ and $\hat{V}[\epsilon]$ as in equations (2.8) and (2.10), respectively.

The resulting approximated interference noise $\hat{\epsilon}_{mj}$ is given by

$$\hat{\epsilon}_{mj} = 10 \log_{10}(1 + z). \quad (2.12)$$

The parameters μ_z and σ_z of the resulting lognormal random variable z are given by

$$\sigma_z = \frac{1}{\lambda} \left(\ln \left(\frac{(\exp(\sigma_{int}^2) - 1) \sum_{k=1}^{M-1} \exp(2\mu_k)}{\left(\sum_{k=1}^{M-1} \exp(\mu_k)\right)^2} + 1 \right) \right)^{\frac{1}{2}} \quad (2.13)$$

and

$$\mu_z = \frac{1}{\lambda} \left(\frac{\sigma_{int}^2 - \sigma_z^2 \lambda^2}{2} + \ln \left(\sum_{k=1}^{M-1} \exp(\mu_k) \right) \right), \quad (2.14)$$

where $\lambda = 10/\ln(10)$ is a constant, and $\mu_k = 10\gamma \log_{10}(\frac{d_{mj}}{d_{kj}})$ and $\sigma_{int} = \sqrt{2}\sigma$ are the parameters of the individual lognormals. Incorporating the values of μ_z and σ_z into equation (2.6), the resulting estimated pdf is given by

$$f_{\hat{\epsilon}}(\hat{\epsilon}) = \frac{10^{\hat{\epsilon}/10}}{\sigma_z \sqrt{2\pi}(10^{\hat{\epsilon}/10} - 1)} \exp \left(\frac{-(10 \log_{10}(10^{\hat{\epsilon}/10} - 1) - \mu_z)^2}{2\sigma_z^2} \right). \quad (2.15)$$

Again, the subscript in $\hat{\epsilon}_{mj}$ has been dropped for clarity. With an expression for $f_{\hat{\epsilon}}(\hat{\epsilon})$, we can now compute the expected interference noise $\hat{\mathcal{E}}[\hat{\epsilon}]$ and variance $\hat{V}[\hat{\epsilon}]$ akin to (2.8) and (2.10) as

$$\hat{\mathcal{E}}[\hat{\epsilon}] = \sum_{n=1}^N \frac{w_n \exp(a_n) a_n 10^{a_n/10}}{\sigma_z \sqrt{2\pi} (10^{a_n/10} - 1)} \exp\left(\frac{-(10 \log_{10}(10^{a_n/10} - 1) - \mu_z)^2}{2\sigma_z^2}\right) \quad (2.16)$$

and

$$\hat{V}[\hat{\epsilon}] = \sum_{n=1}^N \frac{w_n \exp(a_n) (a_n - \hat{\mathcal{E}}[\hat{\epsilon}])^2 10^{a_n/10}}{\sigma_z \sqrt{2\pi} (10^{a_n/10} - 1)} \exp\left(\frac{-(10 \log_{10}(10^{a_n/10} - 1) - \mu_z)^2}{2\sigma_z^2}\right). \quad (2.17)$$

In contrast with (2.8) and (2.10), which are a function of the distance ratio $\frac{d_{1j}}{d_{2j}}$, equations (2.16) and (2.17) are a function of the set of $M - 1$ distance ratios $\frac{d_{mj}}{d_{ij}}$, $i \in [1, 2, \dots, M]$, $i \neq m$. These distance ratios are introduced in the computation of the equivalent lognormal parameter μ_z . Since μ_z is an increasing function of a sum of exponentials of the log-distance-ratios, (2.16) and (2.17) are also increasing functions of the distance ratios (as (2.8) and (2.10) are). This generalizes the behavior observed in the two transmitter case with respect to the distance ratios.

Figure 2.6 shows a comparison of the expressions (2.16) with the result of a Monte Carlo simulation of the interference noise for $M = 3$. As the figure shows, (2.16) approximates the actual expected value of the interference noise within a small error. The figure also shows an increase in absolute error with distance ratio and a maximum error of approximately 0.115 dB. For clarity of presentation, the distance ratios for all interfering transmitters were assumed to be equal. However, similar results are obtained when different distance ratios are observed. As in the two transmitter case, an integration order of $N = 20$ is used.

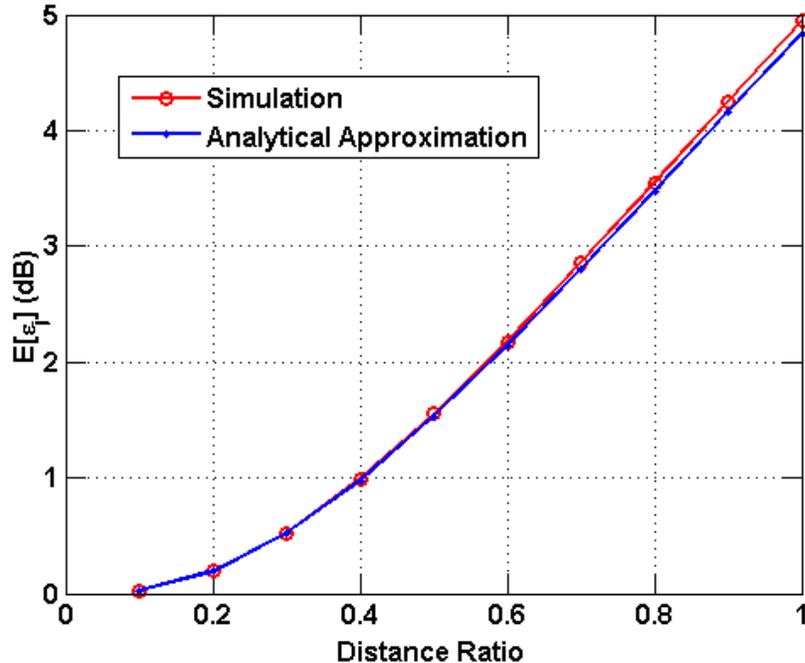


Figure 2.6: Simulated and analytical approximation of $\mathcal{E}[\epsilon]$ for $M = 3$, $\sigma^2 = 36$, and $\gamma = 3$.

2.3 Location Estimation of Multiple Transmitters

In this section we present a technique to estimate the locations of multiple transmitters emitting in the same frequency band. The technique has been named interference subtraction (IS) and is based upon approximating the maximum likelihood (ML) estimate of the transmitter locations when no interference is present. The IS technique segments the multiple transmitter localization problem into independent single transmitter localization tasks by estimating the interference and subtracting it from the observations. The IS methodology can be applied to different estimation techniques and thus offers several ways of estimating the location of multiple transmitters. This trait allows the IS algorithm to offer the flexibility of trading performance for computational complexity and vice versa.

We begin by presenting the ML estimator of the location of a single transmitter. This serves as an introduction since the IS technique is based on approximating this estimator. A subsection on the low-complexity IS algorithm will follow in which we present an

application of the IS methodology to develop a low-complexity estimation algorithm for multiple transmitter localization. We conclude the section by presenting an application of the IS methodology to develop an expectation maximization (EM) algorithm to address the multiple transmitter localization problem. This subsection is preceded by an introduction of the EM algorithm. The two algorithms discussed in this section bring light to the design tradeoff between estimation performance and computational complexity.

2.3.1 Maximum Likelihood Location Estimate of a Single Transmitter

The ML estimate $\hat{\theta}_{ML}$ of the location vector θ of a single transmitter based on the observed power measurements at each cognitive radio node is given by

$$\hat{\theta}_{ML} = \arg \max_{\theta} f(r_{(dB)}|\theta), \quad (2.18)$$

where $f(r_{(dB)}|\theta)$ is the likelihood function of the observations $r_{(dB)}$ conditioned on the transmitter location θ [45]. When only one transmitter is present, the observed power at each node can be expressed as

$$\begin{aligned} r_{j(dB)} &= 10 \log_{10} (P_j \gamma_j) \\ &= 10 \log_{10} (P_j) + X_j. \end{aligned} \quad (2.19)$$

Note that $P_{ij} = P_j$ and $X_{ij} = X_j$ since only one transmitter is active. Since X_j is Gaussian, the received power $r_{j(dB)}$ is also a Gaussian random variable, and thus the likelihood function of the observation vector r_{dB} is given by

$$f(r_{(dB)}|\theta) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(\frac{-(r_{j(dB)} - \mu_j(\theta))^2}{2\sigma_j^2}\right), \quad (2.20)$$

where $\mu_j(\theta) = E[r_j] = 10 \log_{10}(P_j)$ is the expected received power at the j th sensor and is a function of the primary transmitter location θ . Taking the natural logarithm on both sides to obtain the log-likelihood function, letting $\sigma_j = \sigma$ for homogeneous shadowing, and substituting into (2.18), the ML estimate is obtained by solving

$$\hat{\theta}_{ML} = \arg \min_{\theta} \sum_{j=1}^N (r_{j(dB)} - \mu_j(\theta))^2. \quad (2.21)$$

Equation (2.21) represents the result that the IS-based algorithms seek to approximate. This minimization is typically done through search algorithms, since obtaining the minimum through partial derivatives results in a complicated system of equations that is not easily solved.

2.3.2 Low-Complexity Interference Subtraction Algorithm

As mentioned in Section 2.2.2, when multiple transmitters are transmitting simultaneously, the observed received power at each node is given by (2.2), where $\sum_{i=1}^M r_{ij}$ represents a sum of lognormal random variables. Since no analytic distribution exists for the sum of lognormal random variables, no closed-form expression for $P(r_{(dB)}|\theta)$ can be obtained, making computation of $\hat{\theta}_{ML}$ intractable. The IS technique circumvents this problem by approximating the single transmitter case. This is accomplished by estimating the expected interference and subtracting it from the received power measurements at each receiver.

One way of implementing IS is to form sensor clusters that perform independent single transmitter localization. In each cluster, one primary transmitter is assumed to be the main emitter, and the remaining transmitters are treated as interference. Based on the initial location estimates generated by each cluster, the interference power at each receiver is estimated and used to adjust the received power measurements. The adjusted observations are then used to compute an approximated ML estimate of the main emitter for each cluster via (2.18). The procedure for computing the location estimates via this algorithm

is as follows:

1. Cluster receivers into M sets.

2. Generate one initial transmitter location estimate per cluster $\hat{\theta}^{(0)} = [\hat{\theta}_1^{(0)} \ \hat{\theta}_2^{(0)} \ \dots \ \hat{\theta}_M^{(0)}]^T$ according to

$$\hat{\theta}_{io} = \arg \min_{\theta_i} \sum_{j=1}^{K_i} (r_{j(dB)} - E[r_{ij(dB)}|\theta_i])^2, \quad (2.22)$$

where $i = 1, \dots, M$, and K_i is the number of receivers in the i th cluster.

3. For each receiver, compute the adjusted observed power by subtracting an estimate of the received power due to interfering transmitters. The estimated received power at the j th node from the i th transmitter is computed according to

$$\hat{r}_{ij} = r_j - E[\epsilon_{ij}|\hat{\theta}^{(0)}], \quad (2.23)$$

where $E[\epsilon_{ij}|\hat{\theta}^{(0)}]$ is the expected interference noise given the estimated location $\hat{\theta}^{(0)}$.

4. Refine initial location estimates generated in Step 2. The refined estimates $\hat{\theta} = [\hat{\theta}_1 \ \hat{\theta}_2 \ \dots \ \hat{\theta}_M]^T$ are computed according to

$$\hat{\theta}_i = \arg \min_{\theta_i} \sum_{j=1}^K (\hat{r}_{ij(dB)} - E[r_{ij(dB)}|\theta_i])^2, \quad (2.24)$$

where $i = 1, \dots, M$.

5. Choose the final location estimates by evaluating a cost function given by

$$\hat{\theta}_{IS} = \arg \min_{\varphi} (-1) \prod_{j=1}^N \tilde{f}(r_{j(dB)}|\varphi), \quad (2.25)$$

where $\tilde{f}(r_{j(dB)}|\hat{\theta}_{Pi})$ is the approximated likelihood function computed using the Fenton-Wilkinson lognormal approximation [44], and φ can take the values of the 2^M possible combinations of estimates in $\hat{\theta}$ and $\hat{\theta}^{(0)}$ (e.g., for $M = 2$, φ takes the values of the rows of in $[\hat{\theta}_1 \ \hat{\theta}_2 ; \hat{\theta}_1 \ \hat{\theta}_2^{(0)} ; \hat{\theta}_1^{(0)} \ \hat{\theta}_2 ; \hat{\theta}_1^{(0)} \ \hat{\theta}_2^{(0)}]$).

The algorithm begins by forming clusters of receivers. We use the k-means algorithm to perform clustering based on geographic proximity [46]. The k-means algorithm partitions the observations into clusters by iteratively assigning each observation to the cluster with the nearest center of mass. Since interference statistics are closely related to the relative distances between receivers and transmitters, clustering based on proximity helps maintain homogeneous interference statistics across the cluster. Also, because the receivers and transmitters are assumed to be uniformly distributed across the area of interest, proximity-based clustering can be viewed as a first tier of interference rejection, as it increases the likelihood of a cluster sharing the closest transmitter as the one to be estimated. Although k-means was used, any alternative clustering algorithm could potentially be used for this task. Once the receivers are clustered, a first set of estimates $\hat{\theta}^{(0)}$ are computed according to equation (2.22).

In Step 3, the expected interference given $\hat{\theta}^{(0)}$ is estimated. This is done by computing the conditional expectation $E[\epsilon_{ij}|\hat{\theta}^{(0)}]$, which is computed by numerical integration of the expectation integral $\int_0^\infty \epsilon f_\epsilon(\epsilon) d\epsilon$ as discussed in Section 2.2. This estimate is then used to subtract the interference from the observations r_{dB} . In Step 4 of the algorithm, new estimates are computed based on the adjusted observations. At the end of Step 4, we have 2^M possible sets of estimates, and we use the cost function (2.25) to determine which set of location estimates is chosen as the final set of estimates. In contrast to the cost function used for Steps 2 and 4, the cost function of Step 5 is global, i.e., it is a function of all transmitter locations. A global cost function allows us to use an independent relation to obtain relative accuracy without having to minimize a complicated multidimensional function.

The main advantage of the IS algorithm is its low computational complexity compared

to other methods presented in the literature [36]. As our simulation results will show, given computational complexity constraints, the IS algorithm outperforms other approaches by a significant margin. In addition, competing approaches require as much as an order of magnitude increase in computational complexity to obtain a significant increase in performance.

2.3.3 Interference Subtraction and the EM Algorithm

In this subsection we apply the IS methodology to develop an IS-based EM algorithm for multiple transmitter localization, a problem to which the conventional EM algorithm cannot be applied. We begin with a general description of the EM algorithm. The derivation of the IS-based EM algorithm follows.

The EM Algorithm

The EM algorithm is an iterative procedure to approximate the ML estimate of a parameter whose associated realizations are not directly observed. The underlying theory of the algorithm was initially presented independently by several researchers [47, 48]. However, Dempster et al. [49] unified the theory and proved its convergence. A general statement of the algorithm, as presented in [50], is given below.

Let Y denote the sample space of the observations, and let y denote an observation from Y . Let X denote the underlying space and let x be an outcome from X . The data x is referred to as the complete data, and it is not observed directly. It is only observed by means of y , where $y = y(x)$ and $y(x)$ is a many-to-one mapping, i.e. an observation of y determines a subset of X . Given this naming convention, the observation y is also referred to as incomplete data. The pdf of the complete data is $f(x|\theta)$, where θ is a set of parameters of the density function. The pdf of the incomplete data is $g(y|\theta)$, and $L_y(\theta) = \ln g(y|\theta)$ is the log-likelihood function.

It is desired to find θ to maximize the likelihood function $\ln f(x|\theta)$; however, the data needed to compute this function is not available. In this situation, we use the EM algorithm to maximize the expectation of $\ln f(x|\theta)$ given the observed data y and a previously

computed estimate of θ . This two-step process is performed as follows:

1. E-step:

Let $\theta^{[k]}$ be the estimate of the parameters at the k th iteration. Compute the expected likelihood of the complete data given the observation y and the unknown parameter θ

$$Q(\theta|\theta^{[k]}) = \mathcal{E}[L_x(\theta)|y, \theta^{[k]}], \quad (2.26)$$

where $L_x(\theta) = \ln f(x|\theta)$ is the log-likelihood function of the complete data x .

2. M-step:

Compute the value $\theta^{[k+1]}$, which maximizes $Q(\theta|\theta^{[k]})$ as follows:

$$\theta^{[k+1]} = \arg \max_{\theta} Q(\theta|\theta^{[k]}) \quad (2.27)$$

The EM algorithm consists of performing these two steps iteratively until convergence i.e. when $\|\theta^{[k]} - \theta^{[k-1]}\| < \nu$, where ν is chosen to be arbitrarily small or to meet a particular objective. Convergence of the EM algorithm to a local maximum is guaranteed since the algorithm increases the likelihood function at each iteration. However, this maximum is not assured to be a global maximum, and hence the algorithm is often seeded with multiple initial conditions in order to increase the likelihood of convergence to a global maximum. Detailed treatment on the convergence properties of the EM algorithm can be found in [51].

Multiple Transmitter Localization via the EM Algorithm

The problem of multiple transmitter localization via received signal strength does not admit a straightforward application of the EM algorithm, as the complete data cannot be straightforwardly separated or estimated [36]. The observation is a logarithmic function of a sum of lognormal random variable, which prevents the derivation of a closed form expression for $\mathcal{E}[L_x(\theta)|y, \theta^{[k]}]$, complicating the expectation step of the EM algorithm. The IS technique

solves this problem by providing a way of estimating the complete data and facilitating the estimation of $\mathcal{E}[L_x(\theta)|y, \theta^{[k]}]$.

Recall that the received power at the j th cognitive radio node is given by

$$\begin{aligned} r_{j(dB)} &= 10 \log_{10}(r_{mj}) + 10 \log_{10} \left(1 + \sum_{i=1, i \neq m}^M \frac{r_{ij}}{r_{mj}} \right) \\ &= r_{mj(dB)} + \epsilon_{mj}. \end{aligned}$$

The received powers from the desired transmitter $r_{mj(dB)}$, $m \in [1, 2, \dots, M]$, can be defined as hidden variables. The complete data is then given by the matrix

$$\mathbf{H} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_M \end{bmatrix}, \quad (2.28)$$

where the row vectors $h_m = [r_{m1(dB)}, r_{m2(dB)}, \dots, r_{mN(dB)}]$, $m \in [1, 2, \dots, M]$, represent the power received at the cognitive radio nodes from the m th transmitter, i.e., the hidden variables.

Given a current estimate of the primary transmitter locations $\theta^{[k]}$, we use the IS algorithm to estimate the hidden data and construct an expected likelihood function to be maximized. We estimate the hidden data by subtracting the expected interference generated by other (not main) transmitters from the observed data at each receiver. The expected interference is estimated as per Section 2.2, using equation (2.8) or (2.16), depending on the value of

M . We define the $M \times N$ interference matrix as

$$\hat{\mathbf{E}}^{[k]} = \begin{bmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \\ \vdots \\ \hat{\epsilon}_M \end{bmatrix}, \quad (2.29)$$

where the row vectors $\hat{\epsilon}_m = [\mathcal{E}[\epsilon_{m1}|\theta^{[k]}], \mathcal{E}[\epsilon_{m2}|\theta^{[k]}], \dots, \mathcal{E}[\epsilon_{mN}|\theta^{[k]}]]$, $m \in [1, 2, \dots, M]$, contain the estimated interference noise affecting the power measurement from the m th transmitter at each receiver node given the current location estimates $\theta^{[k]}$. We then compute an estimate of the complete data as

$$\hat{\mathbf{H}}^{[k]} = \mathbf{R} - \hat{\mathbf{E}}^{[k]}, \quad (2.30)$$

where \mathbf{R} is an $M \times N$ matrix with each of its rows containing a copy of the observation vector r_{dB} and $\hat{\mathbf{H}}^{[k]}$ is the estimate of the complete data given $\theta^{[k]}$.

The next step is to construct an expected likelihood function of the estimated complete data given the observation and the transmitter locations. We do this by first noting that after performing interference subtraction, each row of $\hat{\mathbf{H}}$, i.e. \hat{h}_m , represents a random vector with the estimated received power from the desired transmitter, which is only a function of a single transmitter location. This allows us to separate the data, compute M expected likelihood functions, and maximize each independently for each transmitter. To determine the likelihood function of the random vectors \hat{h}_m , we note that

$$\begin{aligned} \hat{h}_{mj} &= r_{j(dB)} - \hat{\epsilon}_{mj} \\ &= r_{mj(dB)} + \epsilon_{mj} - \hat{\epsilon}_{mj}, \end{aligned} \quad (2.31)$$

where

$$r_{mj(dB)} = 10 \log_{10}(P_{mj}) + X_{mj} \sim \mathcal{N}(10 \log_{10}(P_{mj}), \sigma^2).$$

The term $\epsilon_{mj} - \hat{\epsilon}_{mj}$ represents the effect of interference subtraction. This term is assumed to be zero-mean with variance σ_ϵ^2 and independent from $r_{mj(dB)}$. As will be shown in Section 2.5.3, the estimate $\hat{\epsilon}_{mj}$ is asymptotically unbiased in the number of sensors, supporting the assumption that $\mathcal{E}[\epsilon_{mj} - \hat{\epsilon}_{mj}] = 0$. Additionally, we model the term $\epsilon_{mj} - \hat{\epsilon}_{mj}$ as a Gaussian random variable that maintains the same mean and variance pair $(0, \sigma_\epsilon^2)$. Figure 2.7 shows a comparison of the actual distribution of the term $\epsilon_{mj} - \hat{\epsilon}_{mj}$ with a Gaussian density having the same mean and variance. Even though the Gaussian density is not an

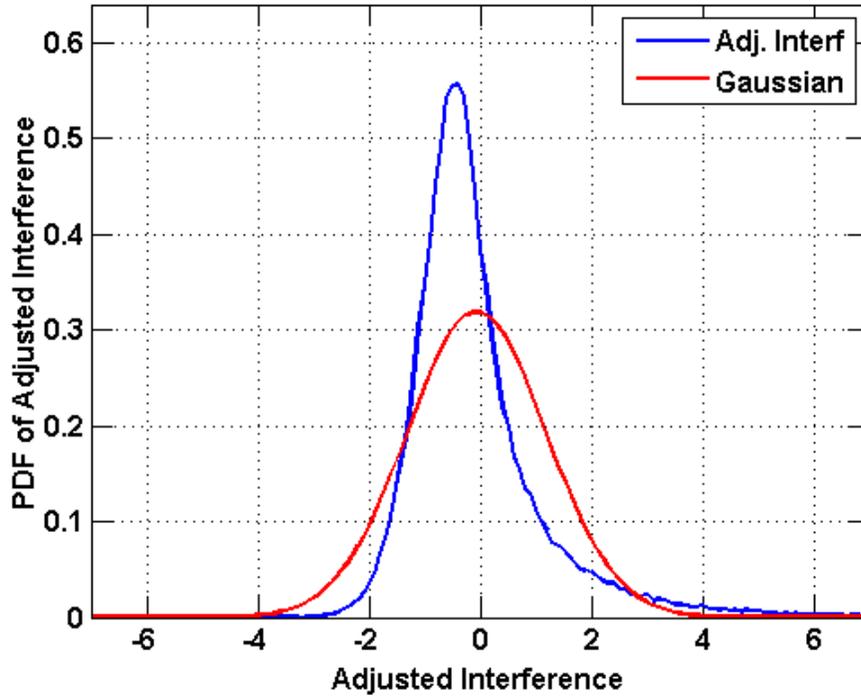


Figure 2.7: Simulated PDF of the adjusted interference term $\epsilon_{mj} - \hat{\epsilon}_{mj}$ vs. Gaussian PDF with same mean and variance, and distance ratio equal to 0.6. $\hat{\epsilon}_{mj}$ was generated given an independent, zero-mean, Gaussian distance ratio estimation error with standard deviation of 0.1.

exact match, it is still employed as it simplifies the analysis while being sufficiently close to provide accurate localization results. Note that the density could be better approximated by

a non-zero-mean Gaussian distribution with a smaller variance. However, it is not pursued, as zero-mean is a desired property that results from interference subtraction. Given this information, we can then determine that the vectors \hat{h}_m are $\mathcal{N}(10 \log_{10}(P_{mj}), \sigma^2 + \sigma_\epsilon^2)$. The expected likelihood of \hat{h}_m can then be computed akin to (2.20) as

$$\mathcal{E}[P(\hat{h}_m|\theta_m)|r_{(dB)}, \theta_m^{[k]}] = \prod_{j=1}^N \frac{1}{\sqrt{2\pi(\sigma^2 + \sigma_\epsilon^2)}} \exp\left(\frac{-(\hat{h}_{mj} - \mu_{mj}(\theta_m))^2}{2(\sigma^2 + \sigma_\epsilon^2)}\right), \quad (2.32)$$

where $\mu_{mj}(\theta_m) = E[r_{mj}|\theta_m] = 10 \log_{10}(P_{mj})$ is the expected received power at the j th sensor conditioned on the m th primary transmitter location θ_m . Taking the natural logarithm and dropping the terms independent of θ_m and constant terms in the denominator, we obtain the expected log-likelihood as

$$\mathcal{E}[L_{\hat{h}_m}(\theta_m)|r_{(dB)}, \theta_m^{[k]}] = \sum_{j=1}^N -(\hat{h}_{mj} - \mu_{mj}(\theta_m))^2. \quad (2.33)$$

Equation (2.33) is equivalent to $Q(\theta|\theta^{[k]})$ of equation (2.26) and is maximized in the M-step of the EM algorithm.

Given the relations derived above, the EM algorithm is performed as follows:

1. Set $k = 0$. Generate initial estimate $\theta^{[k]}$ of the transmitter locations.

2. E-Step:

Compute $\hat{\mathbf{H}}^{[k]}$ given $\theta^{[k]}$ as in equation (2.30).

3. M-Step:

Compute new estimate $\theta_m^{[k+1]} \forall m \in [1, 2, \dots, M]$ by maximizing the the expected

likelihood $\mathcal{E}[L_{\hat{h}_m}(\theta_m)|r_{(dB)}, \theta_m^{[k]}]$ according to

$$\theta_m^{[k+1]} = \arg \min_{\theta_m} \sum_{j=1}^N (\hat{h}_{mj} - \mu_{mj}(\theta_m))^2. \quad (2.34)$$

4. Stop if $\theta^{[k+1]}$ has converged. If not, set $k = k + 1$ and return to step 2.

To increase the likelihood of convergence to a global maximum, the above algorithm can be run for multiple initial conditions in parallel. A cost function such as equation (2.25) could be used to select a final estimate from among the candidates.

2.4 Transmitted Power Estimation

The work performed so far has assumed that the transmitted power from the primary transmitters are known and equal. In practice, this assumption may not be true. In this section we will explore the ramifications of unknown and unequal transmitted powers. We also propose two approaches for jointly estimating transmitted powers and transmitter locations via the EM algorithm.

2.4.1 Incorporating Unknown Transmitted Power to the IS Estimation Framework

Let $P^T = [P_1^T, P_2^T, \dots, P_M^T]$ be a vector containing the transmitted powers from the M primary transmitters, which are assumed unknown and possibly unequal. The power observed at the j th receiver is still given by Equation (2.3). However, its components are now a function of two unknown parameters: P^T and θ .

It is desired to incorporate this new assumption into the interference analysis of Section 2.2 and use the results to derive algorithms to estimate both P^T and θ . We consider the two-transmitter analysis and redefine the interference noise term defined in equation (2.4)

to account for unknown and unequal transmitted powers. This term is now given by

$$\begin{aligned}\epsilon_{1j} &= 10 \log_{10} \left(1 + \frac{r_{2j}}{r_{1j}} \right) \\ &= 10 \log_{10} \left(1 + \frac{P_2^T d_{1j}^\gamma}{P_1^T d_{2j}^\gamma} 10^{\frac{(x_2 - x_1)}{10}} \right),\end{aligned}\quad (2.35)$$

where transmitter 1 is the desired transmitter. The parameters of the distribution of term $1 + \frac{r_{2j}}{r_{1j}}$ inside the logarithm are given by

$$\mu_{int} = 10 \log_{10} \left(\frac{P_2^T d_{1j}^\gamma}{P_1^T d_{2j}^\gamma} \right) \quad \sigma_{int} = \sqrt{2} \sigma_j. \quad (2.36)$$

The pdf of the interference noise can be calculated akin to equation (2.6) and is given by

$$f_\epsilon(\epsilon) = \frac{10^{\epsilon/10}}{\sigma_{int} \sqrt{2\pi} (10^{\epsilon/10} - 1)} \exp \left(-\frac{(10 \log_{10}(10^{\epsilon/10} - 1) - 10 \log_{10}(\frac{P_2^T d_{1j}^\gamma}{P_1^T d_{2j}^\gamma}))^2}{2\sigma_{int}^2} \right). \quad (2.37)$$

Equation (2.37) shows that the pdf of the interference noise ϵ now has an additional parameter, the power ratio $\frac{P_2^T}{P_1^T}$, and hence the statistics of ϵ will depend not only on the node configuration (distance ratio), but on ratio of the transmitted powers as well. Note that while the distance ratio has the interferer term (d_{2j}) in the denominator, the power ratio has the interferer quantity P_2^T in the numerator. This is due to the opposite effect that distance and power have on radio wave propagation.

Given the expression for $f_\epsilon(\epsilon)$, the expected value and variance of ϵ can be approximated using the GL expansion of the expectation integral as in Section 2.2. The approximated

expected value $\hat{\mathcal{E}}[\epsilon]$ is computed as

$$\hat{\mathcal{E}}[\epsilon] = \sum_{n=1}^N \frac{w_n \exp(a_n) a_n 10^{a_n/10}}{\sigma_{int} \sqrt{2\pi} (10^{a_n/10} - 1)} \exp \left(\frac{-(10 \log_{10}(10^{a_n/10} - 1) - 10 \log_{10}(\frac{P_2^T d_{1j}^\gamma}{P_1^T d_{2j}^\gamma}))^2}{2\sigma_{int}^2} \right). \quad (2.38)$$

Similarly, the approximated variance $\hat{V}[\epsilon]$ is computed as

$$\hat{V}[\epsilon] = \sum_{n=1}^N \frac{w_n \exp(a_n) (a_n - \hat{\mathcal{E}}[\epsilon])^2 10^{a_n/10}}{\sigma_{int} \sqrt{2\pi} (10^{a_n/10} - 1)} \exp \left(\frac{-(10 \log_{10}(10^{a_n/10} - 1) - 10 \log_{10}(\frac{P_2^T d_{1j}^\gamma}{P_1^T d_{2j}^\gamma}))^2}{2\sigma_{int}^2} \right). \quad (2.39)$$

Figure 2.8 shows a family of plots of the interference noise statistics vs. power ratio when the distance ratio is fixed to 0.5. Note that the power ratio is plotted up to a value of 2 to allow appreciation of both its positive and negative effects on the interference noise statistics. In comparison with Figure 2.5, the figure shows that $\hat{\mathcal{E}}[\epsilon]$ and $\hat{V}[\epsilon]$ are less sensitive to the power ratio than to the distance ratio. This is expected, as the power ratio is not raised by the path loss exponent as shown in equation (2.36). This is a desired characteristic, as it suggests that estimation errors in transmitted power will have a smaller impact on the accuracy of interference subtraction than will errors in distance ratio.

When $M > 2$, it is straightforward to use the Fenton-Wilkinson equations (2.13) and (2.14) in conjunction with (2.16) to compute $\hat{\mathcal{E}}[\epsilon]$ and $\hat{V}[\epsilon]$ as a function of the power and distance ratios.

2.4.2 Estimating Multiple Transmitter Locations and Transmitted Powers via the EM algorithm

In this section we incorporate the unknown transmitted power into the EM framework developed in Section 2.3.3 and develop an IS-based EM algorithm to estimate P^T and θ . Our goal is to estimate the complete data and obtain a new expected likelihood function

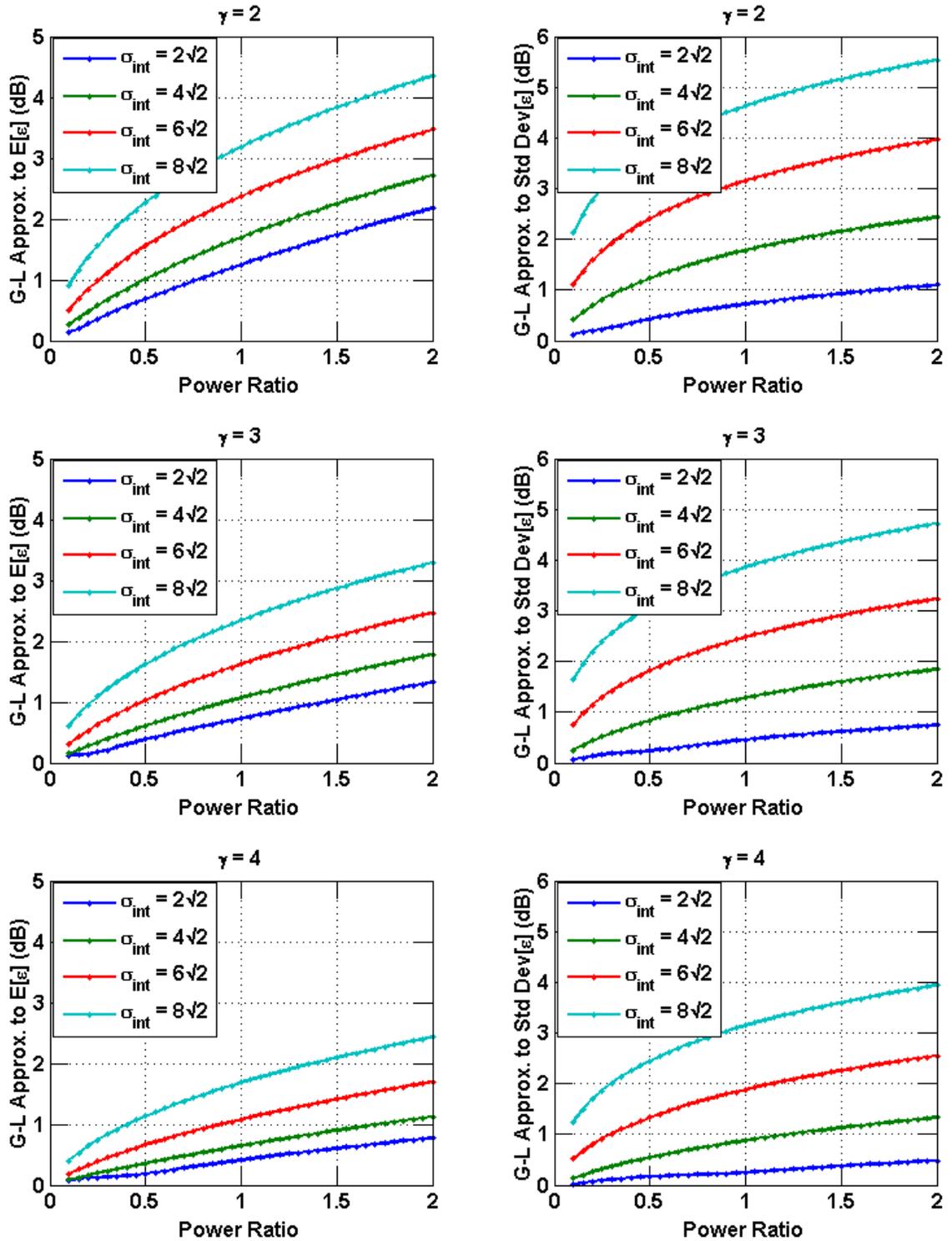


Figure 2.8: Expected value and standard deviation of interference noise vs. power ratio $\frac{P_2^T}{P_1^T}$ for $\frac{d_{1j}}{d_{2j}} = 0.5$, $\gamma = 2$ (top), $\gamma = 3$ (middle), and $\gamma = 4$ (bottom).

$Q(\theta, P^T | \theta^{[k]}, P^{T[k]}) = \mathcal{E}[L_x(\theta, P^T) | y, \theta^{[k]}, P^{T[k]}]$ to be maximized in the M-step of the EM algorithm. To incorporate transmitted power, we can redefine the interference noise matrix defined in (2.29) as

$$\hat{\mathbf{E}}_{\theta, P^T}^{[k]} = \begin{bmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \\ \vdots \\ \hat{\epsilon}_M \end{bmatrix}, \quad (2.40)$$

where the row vectors $\hat{\epsilon}_m = [\mathcal{E}[\epsilon_{m1} | \theta^{[k]}, P^{T[k]}], \mathcal{E}[\epsilon_{m2} | \theta^{[k]}, P^{T[k]}], \dots, \mathcal{E}[\epsilon_{mN} | \theta^{[k]}, P^{T[k]}]]$, $m \in [1, 2, \dots, M]$, contain the estimated interference noise to the m th transmitter at each receiver node given the current location and power estimates $\theta^{[k]}$ and $P^{T[k]}$, respectively. Each interference estimate is obtained by computing its conditional expected value using equation (2.38). The complete data can now be computed akin to (2.30) as

$$\hat{\mathbf{H}}_{\theta, P^T}^{[k]} = \mathbf{R} - \hat{\mathbf{E}}_{\theta, P^T}^{[k]}, \quad (2.41)$$

where $\hat{\mathbf{H}}_{\theta, P^T}^{[k]}$ is the estimate of the complete data given $\theta^{[k]}$ and $P^{T[k]}$. With the complete data estimated, we can now derive the conditional expected likelihood function $Q(\theta, P^T | \theta^{[k]}, P^{T[k]})$.

We carry a similar unbiasedness assumption to that made in Section 2.3.3 and assume that \hat{h}_m is $\mathcal{N}(10 \log_{10}(P_{mj}), \sigma^2 + \sigma_\epsilon^2)$. We then compute the expected log-likelihood akin to (2.33) as

$$\mathcal{E}[L_{\hat{h}_m}(\theta_m, P_m^T) | r_{(dB)}, \theta_m^{[k]}, P_m^{T[k]}] = \sum_{j=1}^N -(\hat{h}_{mj} - \mu_{mj}(\theta_m, P_m^T))^2. \quad (2.42)$$

Equation (2.42) provides a way of jointly estimating P^T and θ . Ideally the estimates would be obtained by taking partial derivatives of these functions with respect to the parameters and solving the system of equations [45]. However, taking the partial derivatives results in

a complicated system of equations that cannot be easily solved through algebraic manipulation. Instead, we use a search algorithm to perform the minimization across the entire parameter space. The addition of parameter P^T represents an increase in the dimensionality of the search space and thus represents a significant increase in computational complexity. We propose a hybrid approach to maximizing the likelihood that avoids adding a dimension to the search. This approach takes advantage of the fact that $\frac{\partial}{\partial P_m^T} \mathcal{E}[L_{\hat{h}_m}(P_m^T)|r_{(dB)}, \theta_m^{[k]}, P_m^{T[k]}]$ can be easily solved for P^T . The approach separates the minimization into two steps: (1) a search over the transmitter locations to find $\theta^{[k]}$, and (2) an evaluation of a closed-form expression to compute $P^{T[k]}$. In the first step, we perform a search to find the locations that maximize

$$\mathcal{E}[L_{\hat{h}_m}(\theta_m)|r_{(dB)}, \theta_m^{[k]}, P_m^{T[k]}] = \sum_{j=1}^N -(\hat{h}_{mj} - \mu_{mj}(\theta_m, P_m^{T[k]}))^2. \quad (2.43)$$

We use the expected log-likelihood function of the known transmitted power case defined in (2.33) conditioned on the current transmitted power estimate $P_m^{T[k]}$. This allows us to estimate the location of the m th transmitter separately from its transmitted power. The second step consists of directly maximizing the expected log-likelihood function of P_m^T conditioned on the current location estimate $\theta_m^{[k]}$ to obtain the transmitted power estimates. This is achieved by taking the partial derivative of the conditional expected log-likelihood as follows:

$$\begin{aligned} \frac{\partial}{\partial P_m^T} \mathcal{E}[L_{\hat{h}_m}(P_m^T)|r_{(dB)}, \theta_m^{[k]}, P_m^{T[k]}] &= \frac{\partial}{\partial P_m^T} \sum_{j=1}^N -(\hat{h}_{mj} - \mu_{mj}(\theta_m^{[k]}, P_m^T))^2 \\ &= \frac{-20\lambda}{P_m^T} \sum_{j=1}^N [\hat{h}_{mj} - P_{m(dB)}^T - 10 \log_{10}(\rho) + 10\gamma \log_{10}(d_{mj}(\theta_m^{[k]}))], \end{aligned} \quad (2.44)$$

where $\lambda = 10/\ln(10)$ and $P_{m(dB)}^T = 10 \log_{10}(P_m^T \rho)$. Equating (2.44) to zero and solving for $P_{m(dB)}^T$ we obtain the following estimate:

$$P_{m(dB)}^{T[k+1]} = -10 \log_{10}(\rho) + \frac{1}{N} \sum_{j=1}^N \hat{h}_{mj} + 10\gamma \log_{10}(d_{mj}(\theta_m^{[k]})), \quad (2.45)$$

where $P_{m(dB)}^{T[k+1]} = 10 \log_{10}(P_m^{T[k+1]})$. Given the above equations, the EM algorithm can be constructed as follows:

1. Set $k = 0$. Generate initial estimates $\theta^{[k]}$ of the transmitter locations and $P^{T[k]}$ of transmitted powers.

2. E-Step:

Compute $\hat{\mathbf{H}}_{\theta, P^T}^{[k]}$ given $\theta^{[k]}$ and $P^{T[k]}$ as in equation (2.30).

3. M-Step:

Compute new estimates $\theta_m^{[k+1]}$ and $P_m^{T[k+1]} \forall d \in [1, 2, \dots, M]$. As discussed above, this step can be performed in two ways. The first is by maximizing the the expected likelihood $\mathcal{E}[L_{\hat{h}_m}(\theta_m, P_m^T) | r_{(dB)}, \theta_m^{[k]}, P^{T[k]}]$ according to

$$[\theta_m^{[k+1]}, P_m^{T[k+1]}] = \arg \min_{\theta_m, P_m^T} \sum_{j=1}^N (\hat{h}_{mj} - \mu_{mj}(\theta_m, P_m^T))^2. \quad (2.46)$$

The second way is to utilize the two-step process developed above to estimate θ_m and P_m^T separately. That is accomplished by maximizing the expected likelihood of equation (2.43) according to

$$\theta_m^{[k+1]} = \arg \min_{\theta_m} \sum_{j=1}^N (\hat{h}_{mj} - \mu_{mj}(\theta_m, P_m^{T[k]}))^2 \quad (2.47)$$

to obtain $\theta_m^{[k+1]}$, and evaluating equation (2.45) to obtain $P_m^{T[k+1]}$.

4. Stop if $\theta^{[k+1]}$ has converged. If not, set $k = k + 1$ and return to step 2.

When the proposed two-part M-step is employed, the EM algorithm presented above represents only a slight computational complexity increase over the known transmitted power case. The increase is mostly due to the evaluation of (2.45) in the M-step, which is done in addition to the 2-D search performed to find the location estimates. This increase in complexity is negligible compared to the 50% increase that an added search dimension would represent [52]. Even though the increase in complexity is small, the overall algorithm complexity is still relatively high compared to the low-complexity IS algorithm described previously.

2.5 IS Performance Analysis

In contrast to estimators for which a closed-form expression can be derived, the performance of the IS technique cannot be characterized using analytical expressions of the estimate statistics. Therefore, we analyze the performance of the IS methodology by looking at an alternative performance measure: the average square error (ASE) of the observation $r_{(dB)}$. We consider three transmitter localization cases and compare the expected ASE to determine the performance differences among the cases. The first of the three cases is the single transmitter case (no interference is present in the measurements). The second case we evaluate is the multiple transmitter case (measurements contain cochannel interference). The third and final case is the multiple transmitter case where the measurements have been adjusted by applying the IS methodology. This formulation allows us to quantify the effects of interference and determine how IS reduces these effects.

2.5.1 Expected Error with No Interference

We begin by looking at the single transmitter case. This serves as a base for comparison throughout our analysis. Let K be the number of cognitive radio nodes whose measurements

are used for estimation, and let

$$e_j = (r_{j(dB)} - \hat{P}_{j(dB)})^2 \quad (2.48)$$

be the square error between the observation at the j th node $r_{j(dB)}$ and the expected observation $\hat{P}_{j(dB)} = 10 \log_{10}(P^T \rho \hat{d}_j^{-\gamma})$ given a transmitter location estimate $\hat{\theta}$. The ASE of the observation, averaged over the receiver nodes, is given by

$$\begin{aligned} E_r &= \frac{1}{K} \sum_{j=1}^K e_j \\ &= \frac{1}{K} \sum_{j=1}^K (r_{j(dB)} - \hat{P}_{j(dB)})^2. \end{aligned} \quad (2.49)$$

Note that equation (2.49) is the same function that is minimized in both the IS and the EM algorithms presented in previous sections. Substituting equation (2.19) into (2.49), expanding the quadratic summand, and taking the joint conditional expectation with respect to the random variables $X_j, j = 1, \dots, K$ conditioned on both the parameter θ and its estimate $\hat{\theta}$, we obtain the following expression for the expected ASE of $r_{(dB)}$:

$$\begin{aligned} \mathcal{E}[E_r | \theta, \hat{\theta}] &= \mathcal{E} \left[\frac{1}{K} \sum_{j=1}^K (P_j + X_j - \hat{P}_j)^2 \right] \\ &= \frac{1}{K} \sum_{j=1}^K (P_j - \hat{P}_j)^2 + \mathcal{E}[X_j^2] \\ &= \frac{1}{K} \sum_{j=1}^K \left(10\gamma \log_{10} \left(1 + \frac{e_{mj}}{d_j} \right) \right)^2 + \mathcal{E}[X_j^2], \end{aligned} \quad (2.50)$$

where $e_{dj} = d_j - \hat{d}_j$ is the error between the estimated and actual distances from the j th receiver to the primary transmitter. Note that the (dB) subscript was removed and all the powers are assumed to be expressed in dB . Equation (2.50) shows that $\mathcal{E}[E_r|\theta, \hat{\theta}]$ is comprised of two terms. The first term relates the distance error e_{dj} to $\mathcal{E}[E_r|\theta, \hat{\theta}]$. This term is a convex function of the distance error e_{dj} with its minimum at $e_{dj} = 0$, i.e., when the $d_j = \hat{d}_j$ the error contribution of this term is zero. The expression also shows an additive shadowing noise squared term, which under homogenous shadowing is constant across all nodes. Equation (2.50) is set as a baseline for comparison as we explore how interference affects $\mathcal{E}[E_r|\theta, \hat{\theta}]$.

2.5.2 Expected Error with Interference

When multiple primary transmitters transmit simultaneously, the observed received power at each node is given by (2.3) and can be separated into three components as follows:

$$r_{j(dB)} = P_{ij} + X_{ij} + \epsilon_{ij}, \quad (2.51)$$

where P_{ij} is the noise-free received power at the j th node from the i th transmitter (the desired transmitter), X_{ij} is the shadowing noise in the channel from the desired transmitter to the j th node, and ϵ_{ij} is the interference noise affecting the measurements of i th transmitter. Substituting (2.51) into (2.49), expanding the quadratic summand, and taking the joint conditional expectations, we obtain the following expression for the expected ASE of

r_{dB} :

$$\begin{aligned}
\mathcal{E}[E_r|\theta, \hat{\theta}] &= \mathcal{E} \left[\frac{1}{K} \sum_{j=1}^K (P_{ij} + X_{ij} + \epsilon_{ij} - \hat{P}_{ij})^2 \right] \\
&= \frac{1}{K} \sum_{j=1}^K (P_{ij} - \hat{P}_{ij})^2 + 2\mathcal{E}[\epsilon_{ij}](P_{ij} - \hat{P}_{ij}) + \mathcal{E}[\epsilon_{ij}^2] + \mathcal{E}[X_{ij}^2] \\
&= \frac{1}{K} \sum_{j=1}^K \left(10\gamma \log_{10} \left(1 + \frac{e_{dj}}{d_{ij}} \right) \right)^2 + \mathcal{E}[\epsilon_{ij}] 20\gamma \log_{10} \left(1 + \frac{e_{dj}}{d_{ij}} \right) \\
&\quad + \mathcal{E}[\epsilon_{ij}^2] + \mathcal{E}[X_{ij}^2]. \tag{2.52}
\end{aligned}$$

Equation (2.52) shows two additional terms in the contribution of each receiver to the total expected error. These two terms are a result of interference noise. The first additional term, which is proportional to the first moment of the interference noise ϵ_{ij} , is a function of the distance error e_{dj} . The addition of this term does not affect the convexity of e_j . However, its minimum is now shifted to a value of $e_{dj} \neq 0$. This shift acts as a bias that distorts the relationship between e_j and the distance error e_{dj} established in (2.50), thus making e_j inaccurate in representing errors in physical distance. The second interference-related term is the second moment of the interference noise. This term does not depend on the distance error, but as was demonstrated in Section 2.2, it varies as a function of the CR node configuration relative to the primary transmitters. This heterogeneous noise can act as an undesired weighting on the error contributions and therefore can affect the minimization result. Equation (2.52) clearly shows the effects of the interference on the expected ASE. Next we will look at how IS reduces these effects.

2.5.3 Expected Error with Interference Subtraction

To analyze the effects of the IS technique on the expected average error of equation (2.52), we are interested in determining how the interference term ϵ_j is changed by the adjustment

applied during the IS procedure. In this section we will define $\tilde{\epsilon}_{ij}$ as the adjusted interference noise and will analyze its statistics to help us achieve this goal.

Let $\hat{\theta}$ be a vector containing estimates of the locations of the main and interfering transmitters, which are computed prior to the IS step. After IS is applied, the adjusted received power at the j th receiver is given by

$$\begin{aligned}\hat{r}_{j(dB)} &= 10 \log_{10}(P_{ij}) + X_{ij} + \epsilon_{ij} - \hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}] \\ &= 10 \log_{10}(P_{ij}) + X_{ij} + \tilde{\epsilon}_{ij},\end{aligned}\tag{2.53}$$

where $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ is computed using equation (2.8), and $\tilde{\epsilon}_{ij} = \epsilon_{ij} - \hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ is the adjusted interference noise. The adjusted interference noise $\tilde{\epsilon}_{ij}$ is assumed to be a zero mean random variable with variance $V[\tilde{\epsilon}_{ij}] = V[\epsilon_{ij}] + V[\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]] - Cov[\epsilon_{ij}, \hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]]$, where $V[\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]]$ is the variance of the estimated expected interference noise $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ and $Cov[\cdot]$ denotes the covariance. The zero mean assumption stems from the fact that the estimate $\hat{\theta}$ is assumed to be an unbiased estimate of θ . Since a closed-form expression for $\mathcal{E}[\hat{\theta}]$ cannot be readily computed, we use simulations to show the validity of this assumption. Table 2.1 shows the simulated average location error of the IS-based EM algorithm presented in Section 2.3.3 over 500 different random draws of $M = 2$ transmitters and $N = 10$ to $N = 40$ receivers with shadowing variance of $\sigma^2 = 36$ and path loss exponent $\gamma = 3$. The results suggest that the location estimates are asymptotically unbiased in the number of receivers, which supports the validity of our assumption when N is sufficiently large. As the table shows, even with N as low as 20 receivers, the average error is small enough to be considered negligible.

To study the expected effect of IS on the observation, we look at the expected adjusted observation. Taking the joint conditional expectation of equation (2.53) over X_{ij} given θ

Table 2.1: Simulated performance of IS-EM algorithm for $M = 2$, $\sigma^2 = 36$ and $\gamma = 3$.

No. of Receivers N	Average Location Error (meters)
10	2.48
20	-0.84
30	0.19
40	0.03

we obtain the expected adjusted receiver observation, which is given by

$$\begin{aligned}\mathcal{E}[\hat{r}_{j(dB)}|\theta] &= 10 \log_{10}(P_{ij}) + \mathcal{E}[X_{ij}] + \mathcal{E}[\tilde{\epsilon}_{ij}|\theta, \hat{\theta}] \\ &= 10 \log_{10}(P_{ij}).\end{aligned}\tag{2.54}$$

Equation (2.54) shows that the expected received power after interference subtraction equals the noise-free power from the main transmitter, suggesting that on the average the IS technique removes the interference noise from the observation.

We now look at the performance metric $\mathcal{E}[E_r|\theta, \hat{\theta}]$. We are particularly interested in the effects of the IS algorithm on (2.52). Substituting the adjusted interference noise $\tilde{\epsilon}_{ij}$ into (2.52), the performance metric becomes

$$\mathcal{E}[E_r|\theta, \hat{\theta}] = \frac{1}{K} \sum_{j=1}^K \left(10\gamma \log_{10} \left(1 + \frac{e_{dj}}{d_{ij}} \right) \right)^2 + \mathcal{E}[\tilde{\epsilon}_{ij}^2] + \mathcal{E}[X_{ij}^2].\tag{2.55}$$

Comparing equation (2.55) with (2.52), we note that the term containing $\mathcal{E}[\epsilon_{ij}]$ has been eliminated. As discussed earlier in this section, this eliminated term distorts the error contribution that each receiver makes to the estimation process. The absence of this term restores the relationship between the estimation error of each receiver e_j and the distance error e_{dj} . We also note that the second interference-related term in (2.55), $\mathcal{E}[\tilde{\epsilon}_{ij}^2]$, is an adjusted version of the term $\mathcal{E}[\epsilon_{ij}^2]$ in (2.52). Since we have assumed that $\mathcal{E}[\tilde{\epsilon}_{ij}] = 0$, this

term can be expanded as follows:

$$\begin{aligned}
\mathcal{E}[\tilde{\epsilon}_{ij}^2] &= V[\tilde{\epsilon}_{ij}] \\
&= V[\epsilon_{ij}] + V[\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]] - Cov[\epsilon_{ij}, \hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]] \\
&= \mathcal{E}[\epsilon_{ij}^2] - \mathcal{E}^2[\epsilon_{ij}] + V[\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]] - Cov[\epsilon_{ij}, \hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]]. \tag{2.56}
\end{aligned}$$

As equation (2.56) shows, this term may be larger than $\mathcal{E}[\epsilon_{ij}^2]$ because of the addition of random variable $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ to the observation in the interference subtraction step. However, as noted previously, since this term is not a function of the distance error, its effects are much less severe than those of the eliminated “distortion” term $\mathcal{E}[\epsilon_{ij}]20\gamma \log_{10} \left(1 + \frac{e_{dj}}{d_{ij}}\right)$. It is interesting to note that if $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ is a consistent estimator, as N approaches infinity, $\mathcal{E}[\tilde{\epsilon}_{ij}^2]$ approaches $\mathcal{E}[\epsilon_{ij}^2] - \mathcal{E}^2[\epsilon_{ij}]$, a value that is guaranteed to be less than $\mathcal{E}[\epsilon_{ij}^2]$.

So far we have looked at the average behavior of the IS algorithm, and we have seen how the interference is removed from the observation and how its effects on the performance metric are reduced. However, the typical error introduced by the application of IS is also of interest. We employ sensitivity analysis to understand how $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ is affected by errors in the estimates of the transmitter locations.

For simplicity, the sensitivity analysis is performed for the two transmitter ($M = 2$) case. The results can be extended to the general ($M > 2$) case by the same methodology presented in Section 2.2. The sensitivity analysis is performed with respect to the distance ratio parameter $\frac{d_{1j}}{d_{2j}}$, the parameter of the distribution of ϵ that is dependent on the location of the transmitters. Using the expression given in (2.8), the sensitivity function of $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ with respect to $\frac{d_{1j}}{d_{2j}}$ can be computed by

$$\frac{\partial}{\partial \frac{d_{1j}}{d_{2j}}} \hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}] = \sum_{n=1}^N \frac{w_n \exp(a_n) a_n \gamma \ln(10) 10^{(a_n/10+1)} d_{2j}}{\sigma_{int}^3 \sqrt{2\pi} (10^{a_n/10} - 1) d_{1j}} \exp\left(\frac{-(K_n)^2}{2\sigma_{int}^2}\right) K_n, \tag{2.57}$$

where

$$K_n = 10 \log_{10}(10^{a_n/10} - 1) - 10\gamma \log_{10} \left(\frac{d_{1j}}{d_{2j}} \right). \quad (2.58)$$

Figure 2.9 shows the sensitivity function vs. distance ratio for various values of γ and σ_{int} . The figure also shows that for most cases the changes in $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ given small errors in $\frac{d_{1j}}{d_{2j}}$ are relatively small compared to the variances of both the shadowing and the interference noise itself. It is worth noting that the sensitivity is given in dB per distance ratio (dB change for an error in distance ratio equal to 1), and since distance ratio errors are most likely to be considerably less than 1, the change in the expected errors in $\hat{\mathcal{E}}[\epsilon_{ij}|\hat{\theta}]$ is expected to be relatively small. The figure also shows that the sensitivity is directly proportional to the shadowing variance, a fact that explains the performance degradation of the IS algorithm as σ increases. Lastly, the sensitivity at low distance ratios (i.e., $\frac{d_{1j}}{d_{2j}} < 0.4$) decreases with γ , an effect that is due to a larger amount of interferer power reaching the receivers when γ is low.

2.6 Simulation Results

In this section we present simulation results for the low-complexity IS and IS-based EM algorithms presented earlier in this chapter. We first describe the simulation setup common to all simulations and then discuss algorithm-specific configurations and results.

2.6.1 Simulation Setup

The simulations were performed for transmitters and receivers lying within a unit square, with certain constraints on the transmitter and receiver geometries. First, transmitters are assumed to be separated by at least 20% of the length of the square, reflecting the physical reality that primary transmitters using the same frequency band would interfere if they were too close together. Second, all receivers are assumed to be at least twice the reference distance d_0 from all transmitters, a constraint required to guarantee that the log-distance

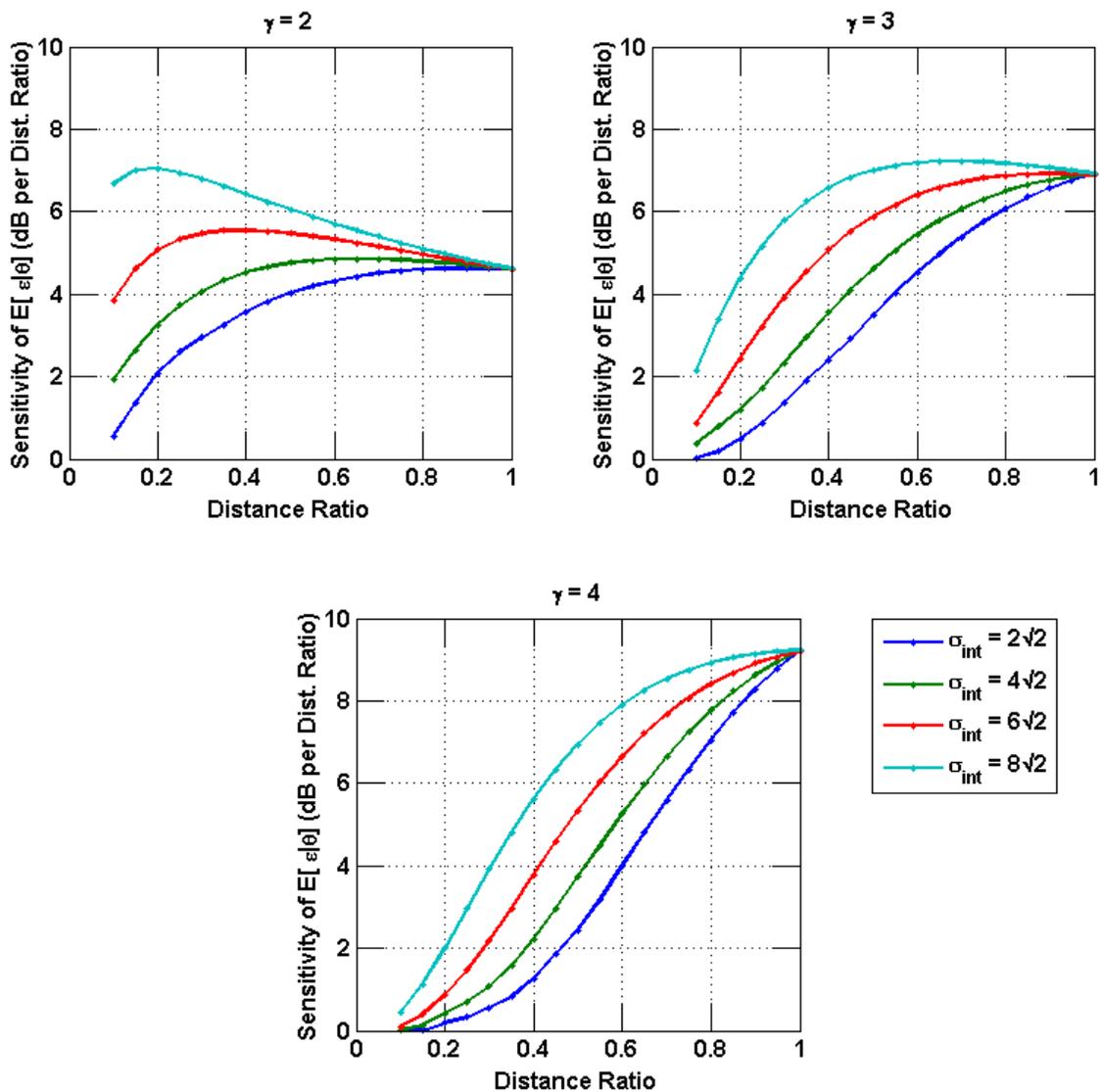


Figure 2.9: Sensitivity of the expected noise with respect to the distance ratio $\frac{d_{1j}}{d_{2j}}$ for $\gamma = 2$ (top left), $\gamma = 3$ (top right), and $\gamma = 4$ (bottom).

propagation model yields realistic results [25].

2.6.2 Low-Complexity IS Algorithm

We compare the performance of the proposed low-complexity IS technique to the quasi expectation maximization (EM) approach proposed in [38]. The quasi EM algorithm is an iterative technique that estimates each transmitter location independently based on an allocated percentage of the power received at each receiver. In the E-step, a percentage of the power received at each receiver is allocated to each transmitter according to the expected received power given the last transmitter location estimates. In the M-step, new estimates are computed based on the allocation obtained on the E-step. To increase the probability of converging to a global minimum, the quasi EM algorithm is run multiple times with different sets of randomly-generated initial conditions.

In order to fairly compare the performance of the low-complexity IS algorithm to that of the quasi EM, we set equivalent constraints on their respective computational complexity levels. Since most of the computation in these two methods is contained in their two-dimensional minimizations, the number of such operations is used as a computational complexity indicator. Both the low-complexity IS algorithm and the quasi EM employ the Nelder-Mead simplex search method to solve their nonlinear minimizations [53, 54]. Given this complexity measure, the computational complexity of the quasi EM is given by $O_{EM} = MIC$, where M is the number of transmitters, I is the number of algorithm iterations, and C is the number of initial estimates. The IS computational complexity is given by $O_{IS} = 2M$, which includes the minimizations performed in steps 2 and 4 of the algorithm. Because the complexity of k-means is negligible compared to that of the Nelder-Mead search method, the computation in Step 1 of the IS algorithm is not included in O_{IS} . In [38] the quasi EM algorithm is run with $C = M^2$ and $I = 10$, a complexity far greater than that of the typical configuration of the IS algorithm. To maintain equal complexity, we run quasi EM with $C = 1$ and $I = 2$. These parameter choices result in an equal number of minimizations regardless of the value of M .

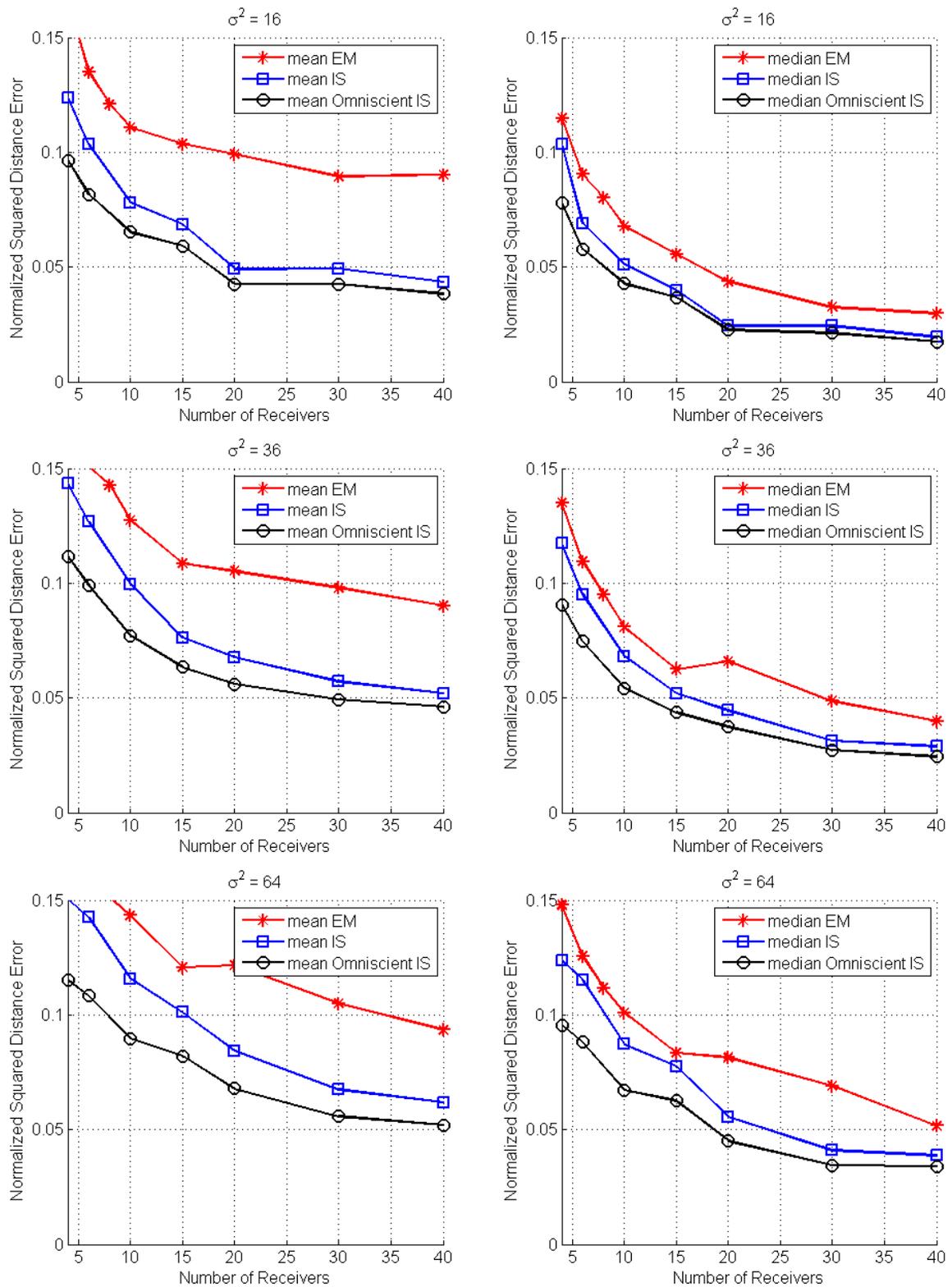


Figure 2.10: Low-complexity IS algorithm two transmitter mean (left column) and median (right column) normalized squared distance error for $\sigma^2 = 16$ (top row), 36 (middle row) and 64 (bottom row).

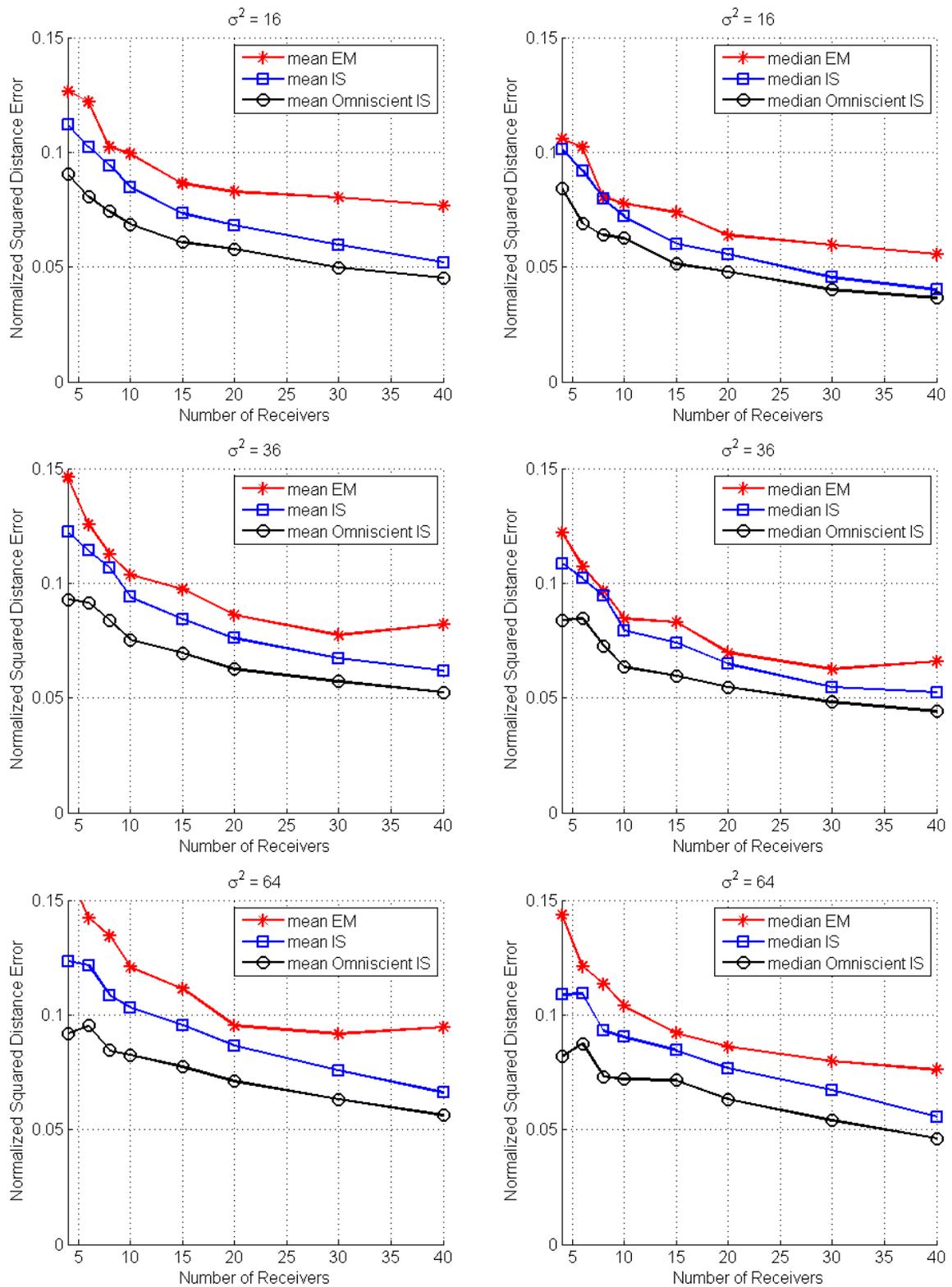


Figure 2.11: Low-complexity IS algorithm three transmitter mean (left column) and median (right column) normalized squared distance error for $\sigma^2 = 16$ (top row), 36 (middle row) and 64 (bottom row).

The simulated performance of the low-complexity IS algorithm and quasi EM is presented in Figures 2.10 and 2.11. The chosen performance metric is the average squared distance error between estimated and true transmitter locations, where the average is taken over the M transmitter location estimates. Performance figures show both the median and mean error over 1000 different random draws of $M = 2$ and 3 transmitters, and $N = 4$ to $N = 40$ receivers with shadowing variance of $\sigma^2 = 16, 36$ and 64. The median error reflects each algorithm's typical performance, while the mean error gives insight into the effects of outlier estimates from each algorithm. Two curves are plotted for the IS algorithm; one is for the algorithm as described in (2.22) through (2.24), and the other is for an omniscient case in which (2.24) is assumed to always provide the estimates with the smallest error. The purpose of showing the omniscient case is to provide an upper bound on the localization accuracy attainable by the IS algorithm. A path loss exponent $\gamma = 2$ was used in all simulations.

As the figures show, the IS approach produces the smallest median error across all values of M , N and σ^2 considered, outperforming the quasi EM in this constrained complexity configuration. This performance gap is larger in mean error, suggesting that the quasi EM produces a larger number of severe localization errors. This gap becomes smaller for larger shadowing variance, most notably in the median error results, indicating that severe shadowing noise poses a significant challenge to both approaches. The performance gap could be attributed to (a) the quasi EM not being able to converge in two iterations; or (b) one initial random guess not being enough for the algorithm to converge to a global minimum. The difference between the standard and the omniscient cases of IS technique increases when $\sigma^2 = 16$, suggesting that the accuracy of (2.24) begins to decrease as σ^2 increases. The slope of both mean and median errors begins to decrease near $N = 15$ but continues to drop as N increases to 40. As the number of receivers grows large, additional power measurements are less likely to provide independent information, and hence the resulting performance improvements are not as significant.

2.6.3 IS-based EM Algorithm

In this subsection we present results for the IS-based EM algorithm developed in Section 2.3.3 and its extension to jointly estimate location and transmitted power discussed in 2.4.2. The IS-based EM algorithm is compared with the low-complexity IS algorithm without any constraints on computational complexity. This allows us to demonstrate the performance-complexity tradeoff inherent in the multiple transmitter localization problem. The problem of jointly estimating location and transmitted power has been treated in the literature [22, 35]. However, it has been done under widely varying assumptions such as having a single transmitter or having information available about the interferer locations. Therefore, results for the IS-based EM algorithm jointly estimating location and transmitted power are presented without comparison to a competing algorithm. The performance of the algorithm is in turn compared to the performance of the “location-only” IS-based EM algorithm to see the effects of the additional parameter on the location estimation performance.

The simulated performance of the IS-based EM algorithm is presented in Figure 2.12. As with the IS algorithm results, the chosen performance metric is the normalized average squared distance error. Performance figures show both the median and mean error over 1000 different random draws of $M = 2$ transmitters and $N = 4$ to $N = 40$ receivers with shadowing variance of $\sigma^2 = 16$ and 36. The IS-based EM algorithm was run with $C = 2$ and $I = 8$. Since the IS-based EM algorithm has equivalent complexity to the quasi EM, its complexity is computed by the same relation, resulting in a complexity of $O_{ISEM} = MIC = 32$. An omniscient case is plotted for each algorithm. Note that in contrast to the quasi EM results presented in the previous section, the IS-based EM algorithm, as run, has an omniscient case because the algorithm is configured with a $C > 1$.

As expected, the IS-based EM algorithm produced the smallest mean and median errors across almost all values of N and σ^2 considered. This increase in performance comes at a cost of an eightfold increase in computational complexity. The performance gap between both algorithms is mostly due to the IS-based EM algorithm’s ability to iterate and reach a better estimate with each iteration. The performance of both algorithms degrades at a

similar pace with an increase in shadowing variance. Additionally, the difference between the actual and omniscient cases is slightly larger for the IS algorithm, suggesting that the EM is capable of more robustly choosing the estimate with the smallest error.

Figure 2.13 shows the performance results for the extended IS-based EM algorithm to jointly estimate location and transmitted power, specifically the version that estimates power and locations separately. The algorithm's performance is compared to that of the non-extended IS-based EM, which only estimates transmitter locations assuming transmitted powers are known. The extended IS-based EM algorithm was configured according to $C = 2$ and $I = 8$ to facilitate comparison with the results of the non-extended version of the algorithm presented above. Performance figure shows both the median and mean errors over 1000 different random draws of $M = 2$ transmitters and $N = 4$ to $N = 40$ receivers with shadowing variance of $\sigma^2 = 16$ and 36 . Both transmitters transmit independent random power levels drawn from a uniform distribution with a minimum of 0.5 Watts and a maximum of 5 Watts. The transmitted power levels are not known by the receivers. As expected, both errors monotonically decrease with increases in N . The figure shows a small decrease in localization performance due to the addition of transmitted power estimation. This is believed to be due to the fact that uncertainty in the transmitted power level adds uncertainty to the observation, causing a decrease in localization accuracy.

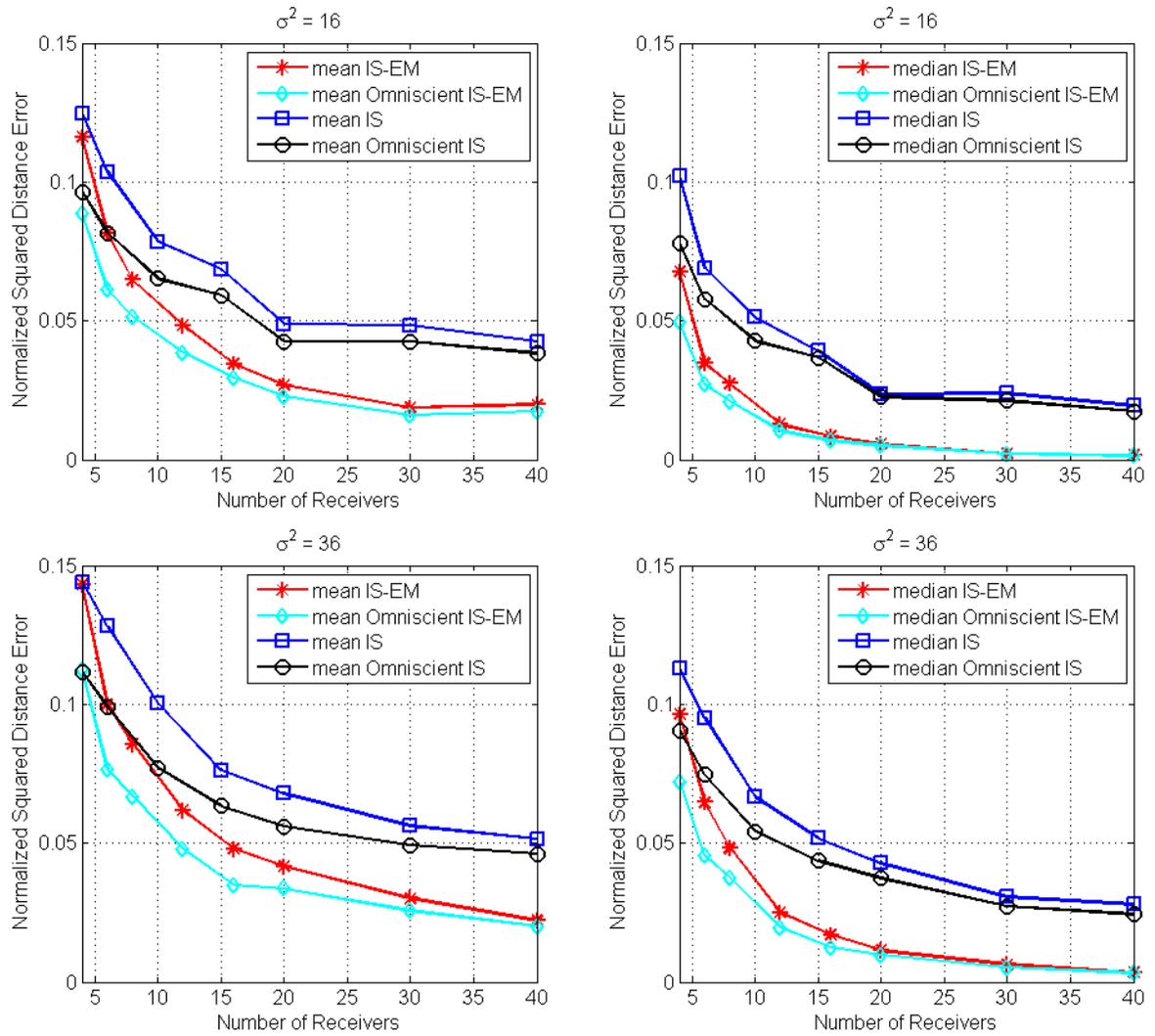


Figure 2.12: IS-based EM algorithm two transmitter mean (left column) and median (right column) normalized squared distance error for $\sigma^2 = 16$ (top row) and 36 (bottom row).

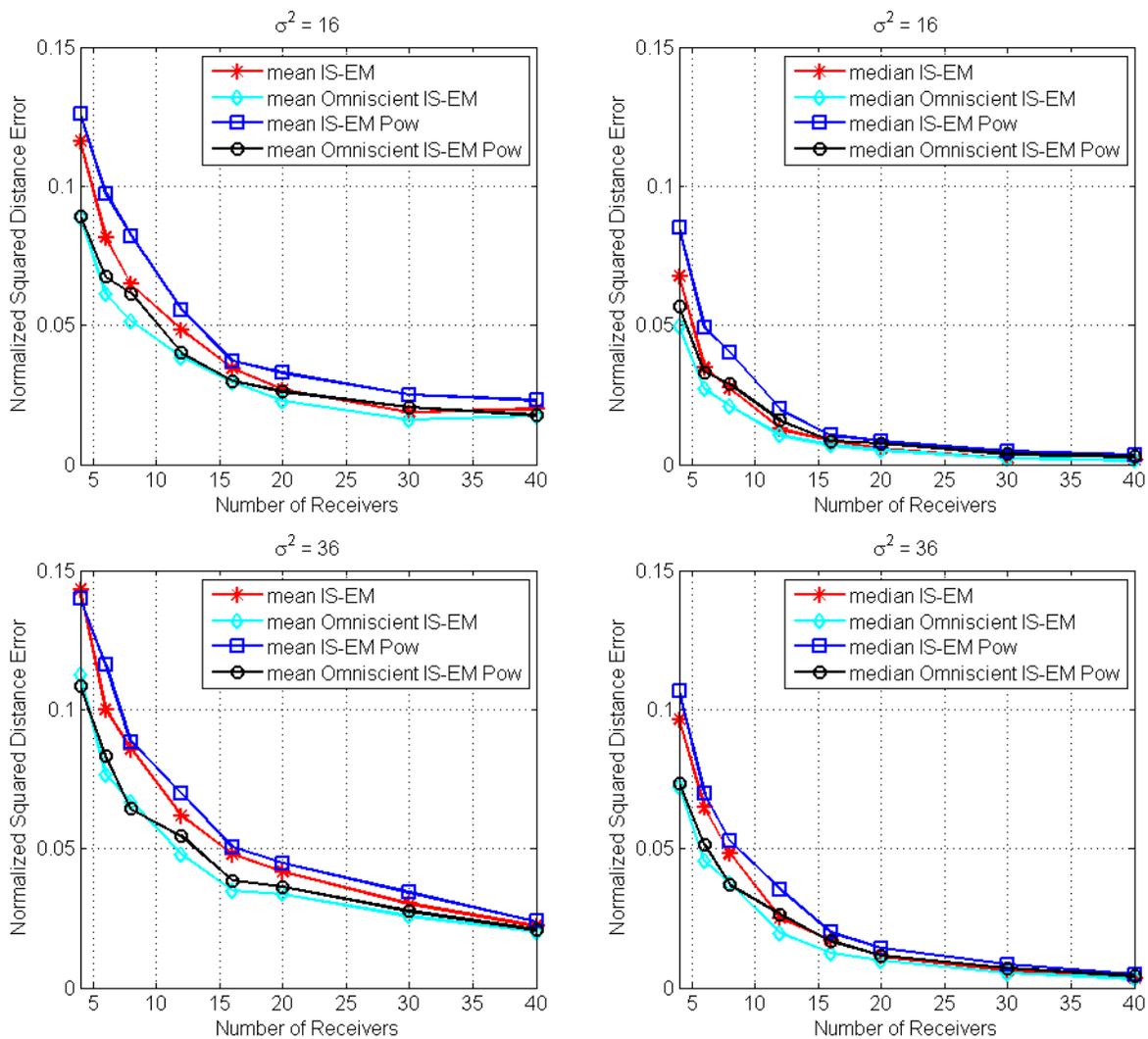


Figure 2.13: Extended IS-based EM algorithm mean (left column) and median (right column) normalized squared distance errors for $\sigma^2 = 16$ (top row) and 36 (bottom row).

Chapter 3: Distributed Processing for Multiple Transmitter Localization

3.1 Background and Related Work

The proposed multiple transmitter localization techniques fall within the realm of distributed estimation since the observations are taken in a distributed fashion. However, the computations for these algorithms are done in a centralized manner, as both the observations and their locations are sent to a centralized fusion center for processing. In applications such as CRNs and WSNs where resources are limited, it may be desirable to spread the computational load uniformly over the network. This can potentially reduce the total energy used for transmitting measurement information. In addition, such approach eliminates the need for a “powerful” node that can act as the fusion center, eliminating single points of failure, and thus, making the system more fault tolerant. To this end, we study the design of distributed algorithms for multiple transmitter localization with the objective of exploring the tradeoffs that energy consumption, computational complexity, and fault tolerance have with localization performance. As is the case with the multiple transmitter localization problem itself, the application of distributed algorithms to the problem has yet to be studied in great depth. Most of the work in this area has concentrated in single source localization for WSN applications where no interference is present in the observations. We will discuss some of the related work throughout this section.

We investigate the problem of distributing localization processing by taking a different approach to multiple transmitter localization. In the proposed approach, we design a localization algorithm specifically for distributed processing. In other words, we investigate algorithms with a specific form such that known distributed processing techniques can be employed for their computations. One example of such approach is the use of a linear

location estimator coupled with a technique referred to as gossiping [55–57] to distribute computation over the network. Gossip algorithms are protocols that dictate how to disseminate and process information throughout the network. These algorithms have been widely studied in computer science for information dissemination and search [58]. Gossip algorithms for distributed processing have been primarily studied as solutions to consensus problems, where a network of agents must achieve a consistent opinion through local information exchanges. We are particularly interested in randomized asynchronous gossip algorithms in which each network node communicates with a randomly selected set of neighbors at randomly selected times [56]. This type of algorithm does not require precise time synchronization nor complicated routing and therefore is well suited for WSN and CRN applications where synchronization and network routing are challenging tasks themselves. Moreover, the randomized gossip algorithm is guaranteed to converge to the equivalent centrally-computed value. A detailed convergence analysis is presented in [56].

A similar approach has been proposed by Rabbat for single acoustic source localization under a Gaussian noise model and has been shown to offer a desirable trade off between performance and computational complexity [59]. In his work, Rabbat proposes an estimator that consists of a linear combination of the receiver locations weighted by a function of their received power. The estimate is given by

$$\hat{\theta} = \frac{\sum_{j=1}^N \phi_j K(r_j)}{\sum_{j=1}^N K(r_j)}, \quad (3.1)$$

where ϕ_j denotes the location of the j th receiver, and K is a monotone increasing function satisfying $K(0) = 0$ and $\lim_{r_j \rightarrow \infty} K(r_j) < \infty$. As mentioned earlier, expressing $\hat{\theta}$ in this manner has the advantage of its computation being readily distributed via algorithms such as the gossip algorithm. Since the work of [59] assumes a single transmitter, the author suggests the choice of the function $K(r_j) = 1_{\{r_j \geq r_{min}\}}$, where $r_{min} > 0$ and $1_{\{\cdot\}}$ is the

indicator function. This reduces (3.1) to

$$\hat{\theta}_1 = \frac{\sum_{j=1}^N \phi_j \mathbf{1}_{\{r_j \geq r_{min}\}}}{\sum_{j=1}^N \mathbf{1}_{\{r_j \geq r_{min}\}}}, \quad (3.2)$$

which represents the average location of the receivers that are within $d_{min} \propto r_{min}^{-1/\gamma}$ from the transmitter. In the absence of noise, this approach draws a circle of radius d_{min} around the transmitter and computes $\hat{\theta}$ using the locations of the receivers inside that circle. In analyzing this approach, we note that it will have difficulties in a cochannel interference environment for two reasons: 1) it relies heavily on the presence of only a single transmitter to select which receivers to use for estimation, and 2) it does not use actual measurement information for the computation of $\hat{\theta}$. There have been similar approaches proposed [60, 61] for single transmitter centralized localization to address 2). However, these approaches, by design, are also susceptible to interference and are not designed with the efficiencies that make them suitable for distributed processing.

As part of our investigation, we apply linear estimators computed distributedly via gossip algorithms to the multiple transmitter localization problem under lognormal shadowing. Specifically, we have developed an iterative distributed clustering method, the multiple transmitter weighted average receiver location (MTWARL), that groups receivers into M clusters in order to apply a linear estimator on each cluster. Our approach uses measurement information to jointly cluster and estimate the transmitter locations. This approach provides increased immunity against cochannel interference and improved accuracy over the mentioned approaches. In what remains of this section, we present the details of the proposed algorithm, perform analysis of the algorithm's convergence properties, energy consumption and computational complexity, and show numerical simulation results of its localization accuracy. The proposed algorithm's energy consumption, computational complexity and localization accuracy are compared to those of the centralized algorithms discussed in Chapter 2, the LCIS and the IS-based EM. Finally, the resulting trade offs are explored.

3.2 Gossip Algorithm

The analysis of the gossip algorithm uses the framework developed by Boyd et al. [56], in which bounds on the convergence time for pairwise randomized gossip algorithms operating over an arbitrary network graph are derived. According to [56], for a particular receiver node configuration, the gossip algorithm to be performed is denoted by

$$\mathcal{A}(P), \tag{3.3}$$

where $P = [P_{mn}]$ is an $N_i \times N_i$ stochastic matrix that dictates the probability that, at a given gossip iteration, the m th and n th nodes will communicate and N_i , $i \in [1, 2, \dots, M]$ is the number of nodes in the i th cluster (nodes that will gossip). Let $x(v) = [\hat{\theta}_{i1}(v), \hat{\theta}_{i2}(v), \dots, \hat{\theta}_{iN_i}(v)]$ be a vector containing the i th transmitter estimate at each receiver during the v th gossip iteration. During each gossip iteration, this value is updated according to

$$x(v) = W(v)x(v-1), \tag{3.4}$$

where $W(v)$ is a random mixing matrix. The value $W(v)$ depends on which nodes are chosen to communicate during a gossip iteration. Thus, with probability $\frac{1}{N_i}P_{nm}$, $W(v)$ is

$$W_{nm} = I - \frac{(e_n - e_m)(e_n - e_m)^T}{2}, \tag{3.5}$$

where e_n is an $N_i \times 1$ unit vector with the n th component equal to 1. The random matrix $W(v)$ essentially averages (in place) the n th and m th estimates in $x(v)$. As it will be shown in Section 3.4, the properties of this matrix govern the convergence behavior of the algorithm.

3.3 Distributed MTL via the MTWARL Algorithm

In this section, the MTWARL algorithm for distributed multiple transmitter localization is detailed. The MTWARL algorithm consists of a nested loop in which the outer loop indexes the transmitters to be localized, and the inner loop computes and refines the location estimates. The estimator used within the algorithm is the average receiver location weighted by r_j , which results from equation (3.1) with a weighting function $K(r_j) = r_j$. This approach weights more heavily the locations of nodes likely to be close to a transmitter, thus reducing the influence of measurements at more distant sensors, which are likely to suffer from higher interference. Pairwise randomized asynchronous gossiping is employed for distribution of the computations over the network [56]. Under this type of gossip algorithm, randomly selected pairs of nodes iteratively share and update their estimates. The number of gossip iterations is determined based on an allowable convergence error. This topic will be discussed in detail in Section 3.4.2.

The MTWARL algorithm is implemented as follows. Let the space $\mathcal{S} = \{j\}_{j=1}^N$ be the set containing the indices of all receivers and $\mathcal{C}_0 = \{\}$ be the empty set. Set the inner iteration count k to zero. Then, for all $i \in [1, 2, \dots, M]$ perform the following procedure:

1. Define the initial set of receivers used to estimate the location of transmitter i according to one of two possible methods. These possible initial clusters are defined as

$$\mathcal{C}_i = \left(\bigcup_{z=0}^{i-1} \mathcal{C}_z \right)', \quad (3.6)$$

where $()'$ denotes the complement of a set; or as

$$\mathcal{C}_i = \{j : d_{xj} < d_{cl}\}, j \in [1, 2, \dots, N], x \neq j, \quad (3.7)$$

where d_{xj} is the distance between the j th receiver and the receiver with the largest power measurement (with index denoted by $j = x$), and d_{cl} is the clustering radius.

The clustering radius is chosen as a fraction of the area of interest and depends on the number of transmitters present and on the propagation environment. Compute an initial estimate of the location of the i th transmitter as

$$\hat{\theta}_i^k = \frac{\sum_{j \in \mathcal{C}_i} \phi_j r_j}{\sum_{j \in \mathcal{C}_i} r_j}. \quad (3.8)$$

2. Cluster by redefining the set \mathcal{C}_i as follows:

$$\mathcal{C}_i = \left\{ j : d_{ij}^k < d_{cl} \right\}, \quad j \in [1, 2, \dots, N], \quad (3.9)$$

where d_{ij}^k is the distance between the j th sensor and $\hat{\theta}_i^k$.

3. Update the estimate of θ_i using the receivers in \mathcal{C}_i by computing

$$\hat{\theta}_i^{k+1} = \frac{\sum_{j \in \mathcal{C}_i} \phi_j r_j}{\sum_{j \in \mathcal{C}_i} r_j}. \quad (3.10)$$

4. Check for convergence using the following condition:

$$\left\| \hat{\theta}_i^k - \hat{\theta}_i^{k+1} \right\| < \nu, \quad (3.11)$$

where ν is an arbitrarily chosen small distance. If condition is met, increment i , set k to zero, and go to step 1. If condition is not met, increment k and go to step 2. Alternatively, the algorithm may be configured to run a fixed number of inner loop iterations.

As the above enumeration shows, for each transmitter, the algorithm starts by forming an initial cluster that will be tasked with estimating the location of a single transmitter. This cluster is formed based on proximity to a particular location. The simplest way that

this is achieved is to use (3.6) to include all receivers located across the entire area of interest. As mentioned above, to save transmit energy, receivers can be grouped around the receiver with the maximum received power measurement. This technique involves disseminating the measurements among receivers in order to determine the largest, thus it does not come for free. However, the energy utilized for this is much less than gossiping over the entire area interest. Energy utilization can be further reduced by, on the first outer loop iteration ($i = 1$), clustering around a smaller area centered at the center of the area of interest. This option saves dissemination of the measurements over the largest possible area, with little effect to localization accuracy. For $i > 1$, the maximum power can be determined distributedly without using excessive energy, since the dissemination area decreases with i . Once the initial cluster has been formed, two parallel instances of the randomized gossip algorithm are spawned over the cluster to compute the numerator and denominators of equation (3.8). These two instances share the same transmissions, but are updated and kept separately at each node. After the gossiping is completed, the cluster \mathcal{C}_i is updated according to (3.9), where the new cluster center is the newest location estimate $\hat{\theta}_i^k$. Nodes outside of the updated cluster would stop gossiping, while new members join in when contacted by another node requesting to gossip. A new instance of the gossip algorithm is then spawned over this newly updated cluster. This process, which continues until (3.11) is satisfied, is repeated for each i . At the end of the final inner loop iteration for each transmitter, the nodes at the edge of each cluster broadcast their result to neighboring nodes, which can then disseminate the result over their respective clusters.

The MTWARL offers a simple, yet robust, method for distributed multiple transmitter localization. To better understand its viability, analysis of several key system parameters such as energy consumption and computational complexity needs to be performed. In the following sections these are parameters are explored with the objective of comparing them to those of the centralized multiple transmitter localization algorithms.

3.4 Gossip Algorithm Convergence

In this section, the convergence properties of the gossip algorithm as they relate to the proposed distributed algorithm, the MTWARL, are examined. As discussed in section 3.3, the MTWARL employs pairwise randomized asynchronous gossiping for distribution of its computation over the network. Under this type of gossip algorithm, at each iteration, a randomly selected pair of nodes share and update their estimates. This process is repeated until all receivers are likely to have converged (within some error) to the equivalent centralized location estimate. The number of iterations required for convergence, called convergence time, influences key system parameters such as energy consumption and computational complexity. Therefore, it is important to study the convergence time of the gossip algorithm as applied for multiple transmitter localization to ensure that the algorithm is not only feasible, but that it also operates as efficiently as possible. To achieve this, two key analysis are performed: 1) convergence time bounds are studied and the optimum gossip configuration that yields the lowest convergence time bounds is derived, and 2) the effects of convergence errors on localization accuracy are explored and stop criteria for energy efficient gossiping are determined.

Given a gossip algorithm $\mathcal{A}(P)$, the ψ -convergence time is defined as

$$T(\psi, P) = \sup_{x^{[0]}} \inf \left\{ v : \Pr \left(\frac{\|x(v) - \hat{\theta}_i^c \mathbf{1}\|}{\|x(0)\|} \geq \psi \right) \leq \psi \right\}, \quad (3.12)$$

where $\mathbf{1}$ is a vector of size N with all ones, $\|\cdot\|$ is the l_2 norm, and $\hat{\theta}_i^c$ is the centralized equivalent estimate of the i th transmitter as defined in (3.8). Thus, $T(\psi, P)$ is the smallest time it takes for $x(\cdot)$ to get within ψ of $\hat{\theta}_i^c \mathbf{1}$ with high probability. Based on the statistical properties of $W(v)$, Boyd proves that $x(v)$ converges to $\hat{\theta}_i^c$ in expectation and also derives

bounds on the convergence time $T(\psi, P)$. These bound are given as

$$\frac{0.5 \log(\psi^{-1})}{\log(\lambda_2^{-1})} \leq T(\psi, P) \leq \frac{3 \log(\psi^{-1})}{\log(\lambda_2^{-1})}, \quad (3.13)$$

where λ_2 is the second largest eigenvalue of the expected value of the mixing matrix $W(v)$.

The expected value of $W(v)$ is given by

$$\mathcal{E}[W(v)] = \frac{1}{N_i} \sum_{n,m} P_{nm} W_{nm} \quad (3.14)$$

$$= I - \frac{1}{2N_i} D + \frac{P + P^T}{2N_i}, \quad (3.15)$$

where $D = \text{diag}([D_1, \dots, D_{N_i}])$ is the diagonal matrix with diagonal entries

$$D_n = \sum_{m=1}^{N_i} [P_{nm} + P_{mn}]. \quad (3.16)$$

As (3.16) shows, matrix D contains the sum of the transpose elements of P that compose the diagonal of matrix $\mathcal{E}[W(v)]$ as part of the expectation summations in (3.14).

3.4.1 Optimum Gossip Algorithm Configuration

As seen in equation (3.13), the convergence time are monotonically increasing function of λ_2 . Therefore, it is of interest to find the matrix P such that λ_2 is the smallest, while satisfying constraints on P . Given this information, the optimum matrix P^* can be found by solving the following optimization problem:

$$\begin{aligned} & \text{minimize} && \lambda_2 \\ & \text{subject to} && \mathcal{E}[W(v)] = I - \frac{1}{2N_i} D + \frac{P+P^T}{2N_i} \\ & && P_{nm} \geq 0, P_{nn} = 0 \end{aligned}$$

$$\sum_m P_{nm} = 1, \forall i.$$

The objective function, which is the second largest eigenvalue of the doubly stochastic matrix $\mathcal{E}[W(v)]$, is a convex function on the set of symmetric matrices P . Therefore, the above problem is a convex optimization problem. Since the objective function requires the computation of all the eigenvalues of $\mathcal{E}[W(v)]$ and then selecting the second largest, optimizing it may be inefficient and difficult to implement using standard convex optimizers. In view of this, it is desirable to find a closed-form expression for λ_2 that can be readily optimized. Observing that $\mathcal{E}[W(v)]$ is a stochastic matrix, it is known that $\lambda_1 = 1$ and that it is associated with the eigenvector $\mathbf{1}$. Therefore, λ_2 can be expressed as the norm of the matrix $\mathcal{E}[W(v)]$, restricted to the subspace $\mathbf{1}^\perp$. Taking advantage of this property and the fact that $P\mathbf{1} = \mathbf{1}$, a closed-form expression for λ_2 is given as

$$\begin{aligned} \lambda_2 &= \|Q\mathcal{E}[W(v)]Q\| & (3.17) \\ &= \|(I - (1/N_i)\mathbf{1}\mathbf{1}^T)\mathcal{E}[W(v)](I - (1/N_i)\mathbf{1}\mathbf{1}^T)\| \\ &= \|(\mathcal{E}[W(v)] - (1/N_i)\mathbf{1}\mathbf{1}^T)\|, \end{aligned}$$

where $Q = (I - (1/N_i)\mathbf{1}\mathbf{1}^T)$ is the matrix representing the orthogonal projection on $\mathbf{1}^\perp$.

Given (3.17), the optimization problem to obtain P^* can be restated as

$$\begin{aligned} \text{minimize} \quad & \|(\mathcal{E}[W(v)] - (1/n)\mathbf{1}\mathbf{1}^T)\| \\ \text{subject to} \quad & \mathcal{E}[W(v)] = I - \frac{1}{2N_i}D + \frac{P+P^T}{2N_i} \\ & P_{nm} \geq 0, P_{nn} = 0 \\ & \sum_m P_{nm} = 1, \forall i. \end{aligned}$$

The above problem can be readily solved by an optimization package such as CVX [62, 63], which converts the problem into a semidefinite program and then uses interior point methods [64] to efficiently solve it. Using such technique, the optimization problem yields

the following result:

$$\begin{aligned}
 P^* &= \frac{1}{(N_i - 1)} (\mathbf{1}\mathbf{1}^T - I) \tag{3.18} \\
 &= \begin{bmatrix} 0 & 1/(N_i - 1) & \cdots & 1/(N_i - 1) \\ 1/(N_i - 1) & 0 & \cdots & 1/(N_i - 1) \\ \vdots & \vdots & \ddots & \vdots \\ 1/(N_i - 1) & 1/(N_i - 1) & \cdots & 0 \end{bmatrix}.
 \end{aligned}$$

As equation (3.18) shows, the optimum stochastic matrix for the case of a fully connected cluster of nodes (i.e., a complete graph) consists of the setup where all pairs of nodes are equally likely to communicate at any given gossip iteration. This resulting configuration is very important as it is optimum with respect to convergence time. Since convergence time and energy consumption are directly proportional, this optimum configuration helps bring down energy consumption by ensuring that, given a desired convergence threshold ψ , only the minimum amount data is transmitted. The energy consumption of the proposed MTWARL algorithm is analyzed in a subsequent section.

Given the optimum matrix P^* defined in (3.18), the bounds on convergence time for this optimum configuration can be computed using equation (3.13). Figure 3.1 shows these bounds for $\psi = 0.05$ compared to the simulated number of iterations that yields the same value of ψ . Simulations were performed under the same setup described in Section 2.6 using P^* as the stochastic matrix for the gossip algorithm. In addition, the number of gossip iterations or stop criteria was set according to the linear function

$$I_g = F_g N_i, \tag{3.19}$$

where F_g is the gossip iteration factor. This stop criteria was chosen to be a linear function of N_i to match the order of the bounds (3.13), which are also a linear function of N_i . A

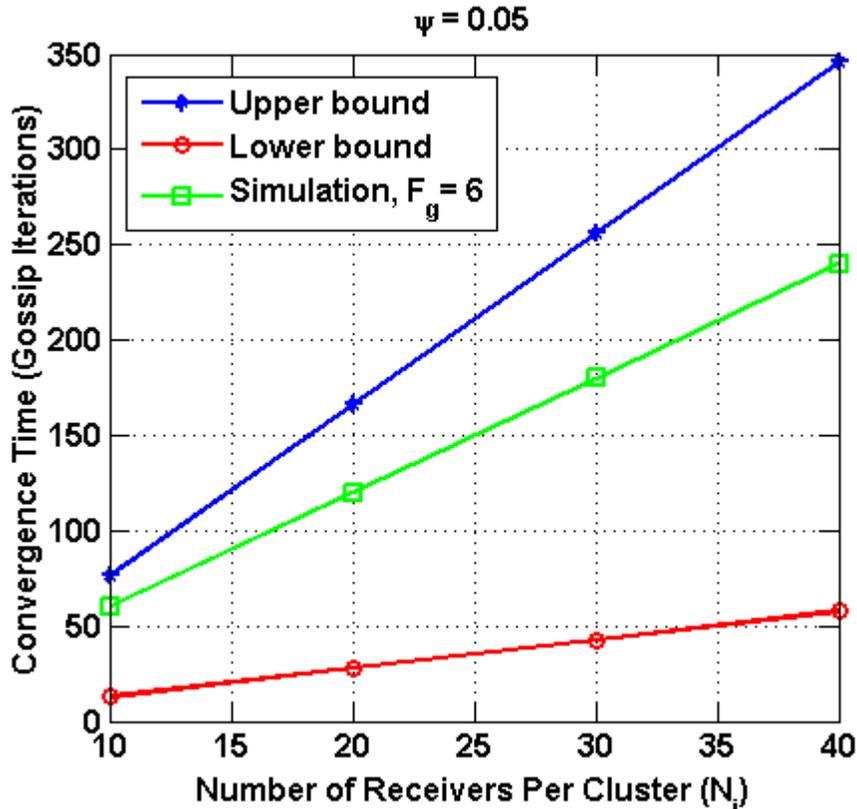


Figure 3.1: Bounds on convergence time for $\psi = 0.05$ vs. number of receivers compared to the simulated number of iterations that yields the same ψ value.

value of $F_g = 6$ was used for the simulation. As the plot shows, the number of gossip iterations used for the simulation lied within the bounds on ψ -convergence time established in equation 3.13.

3.4.2 Convergence Effects on Localization Performance

As discussed in prior sections, energy consumption is directly proportional to convergence time. Therefore, to minimize energy consumption, it is of interest to find the smallest convergence time that results in negligible loss of localization accuracy. In other words, it is of interest to find the largest allowable convergence error threshold ψ that yields no considerable loss in localization accuracy. In this section, we explore the effects of convergence errors on localization performance and develop a framework for expressing

localization error due to gossip convergence errors. Since differences in form between the measures of convergence and localization errors prevent expressing one as a function of the other, numerical simulations are utilized to relate these values and to establish an appropriate gossip algorithm stop criteria for energy efficiency.

The convergence error as defined in (3.12) is given by

$$E_t = \frac{\|x(v) - \hat{\theta}_i^c \mathbf{1}\|}{\|x(0)\|}. \quad (3.20)$$

This error expression is suitable for convergence analysis as it measures the normalized "distance" between the estimate computed distributedly and the estimate computed centrally. As (3.20) shows, the convergence error E_t approaches 0 as the vector of estimate values at each receiver approaches the centrally computed estimate. As discussed in Section 2.6, the normalized localization error due to the random noise contained in the measurements is given by

$$\begin{aligned} E_\theta &= \frac{1}{M} \sum_{i=1}^M \|e_i^\theta\|^2 \\ &= \frac{1}{M} \sum_{i=1}^M \|\hat{\theta}_i^c - \theta_i\|^2. \end{aligned} \quad (3.21)$$

In contrast to (3.20), the error in (3.21) measures distance squared between the estimated and the actual transmitter locations. It is desirable to express the errors due to convergence as a distance squared quantity, so that it can be modeled as an additive quantity within (3.21). Let

$$e_i^t = \hat{\theta}_i^d - \hat{\theta}_i^c \quad (3.22)$$

be the two-dimensional localization error due to convergence errors affecting the i th transmitter estimate, where $\hat{\theta}_i^d$ is the location estimate for the i th transmitter computed distributedly. Solving (3.22) for $\hat{\theta}_i^d$ and substituting it into (3.21), the total distributed localization error can be obtained as

$$\begin{aligned}
 E_\theta^d &= \frac{1}{M} \sum_{i=1}^M \left\| \hat{\theta}_i^d - \theta_i \right\|^2 \\
 &= \frac{1}{M} \sum_{i=1}^M \left\| e_i^\theta \right\|^2 + \left\| e_i^t \right\|^2 + 2e_i^t \cdot e_i^\theta \\
 &= E_\theta + E_d,
 \end{aligned} \tag{3.23}$$

where $E_d = \frac{1}{M} \sum_{i=1}^M \left\| e_i^t \right\|^2 + 2e_i^t \cdot e_i^\theta$ is the normalized localization error due to gossip convergence errors. To put the losses in accuracy due to convergence errors in greater perspective, (3.23) can be solved for E_d and normalized by E_θ to obtain the accuracy loss as a fraction of the localization error due to random factors E_θ . This expression is given by

$$\begin{aligned}
 L_d &= \frac{E_d}{E_\theta} \\
 &= \frac{E_\theta^d - E_\theta}{E_\theta}.
 \end{aligned} \tag{3.24}$$

In order to determine the effects of convergence errors on localization accuracy, it is desirable to link the convergence error E_t with the E_d . However, since E_t is dependent on the initial conditions, it is impossible to express E_d as a function of E_t . To circumvent this problem, numerical simulations are utilized to explore this connection and determine appropriate number of gossip iterations for energy efficient operation. Since the interest lies in the difference between centralized and distributed localization results, an observable that is independent of the propagation conditions and number of transmitters, evaluating a single propagation environment (fixed values of γ and σ) containing two transmitters is

sufficient to draw conclusive results applicable to all localization scenarios. Moreover, these results can be generalized to the application of pair-wise randomized gossip algorithm to any transmitter localization algorithm.

Numerical simulations were performed under the same setup described in Section 2.6. The gossip algorithm was setup with the optimum stochastic matrix P^* and the number of gossip iterations I_g was set according to equation (3.19). Figure 3.2 shows the mean accuracy loss due to gossip convergence errors L_d for different values of F_g and the corresponding convergence error threshold ψ . The figure shows the mean loss over 1000 different random

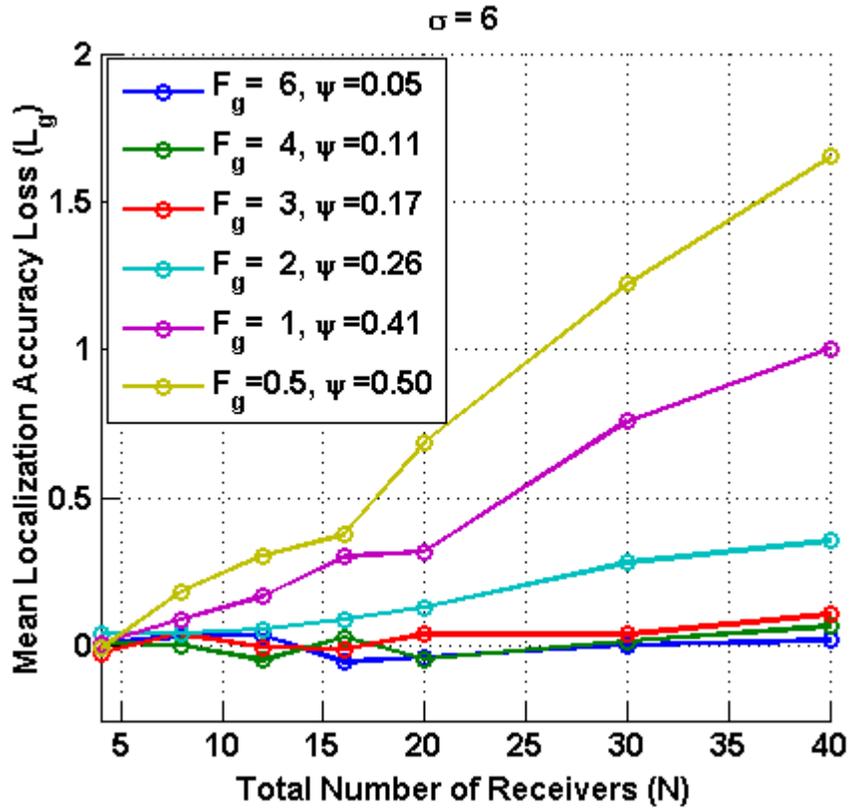


Figure 3.2: Mean localization accuracy loss due to gossip convergence errors for different values of the gossip iteration factor F_g and $\sigma^2 = 36$.

draws of $M = 2$ transmitters and $N = 4$ to $N = 40$ receivers under a shadowing variance of $\sigma^2 = 36$. As seen in the figure, the mean accuracy loss increases with N . This is believed to be due to the E_θ^d being relatively constant with N and thus becoming an increasing

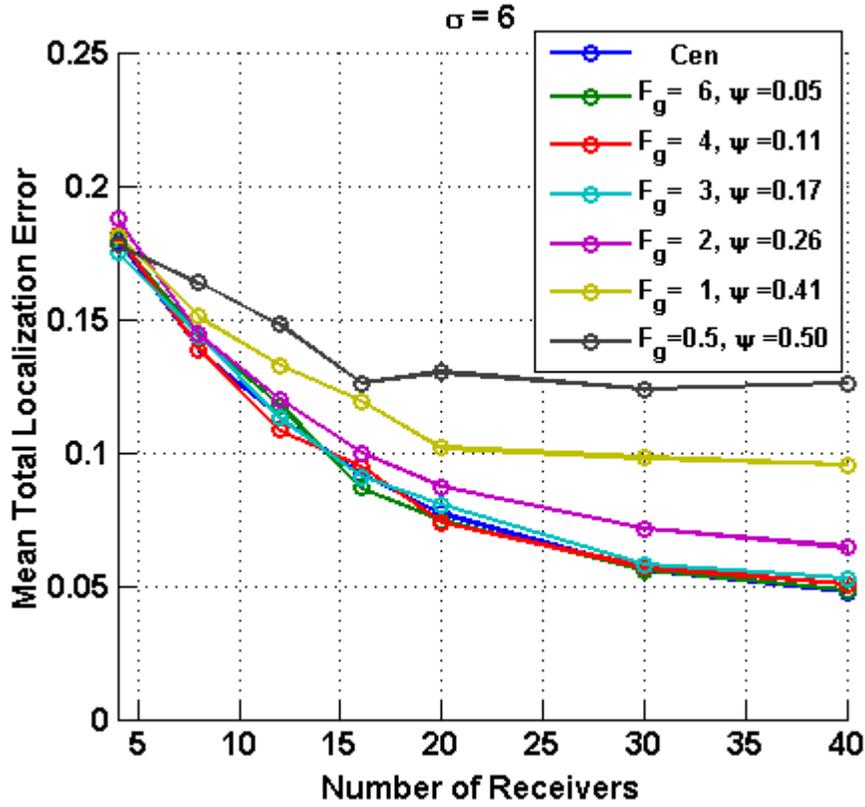


Figure 3.3: Two-transmitter mean localization performance of the MTWARL algorithm for different values of the gossip iteration factor F_g and $\sigma^2 = 36$.

fraction of E_θ as it decreases. This is more evident at values of F_g smaller than 3, where E_θ^d is on the same order as E_θ . The figure also shows that L_g is within $\pm 10\%$ of E_d for values of F_g greater or equal to 3. The figure also shows accuracy losses significantly greater than 10% for values of F_g greater than 3. This is a very important result, which suggests that the gossip algorithm will converge to a low enough convergence error to not degrade accuracy with as little as $I_g = 3N_i$ gossip iterations. Figure 3.3 shows the corresponding total localization accuracy versus total number of receiver plot. Both figures show the values of ψ that correspond to each gossip iteration factor. These values were computed in reverse of the definition of ψ -convergence time (3.12), where instead of finding the number of iterations that met the convergence criteria for a given ψ , the value of ψ was determined given a fixed number of iterations. The values of ψ were determined by constructing the

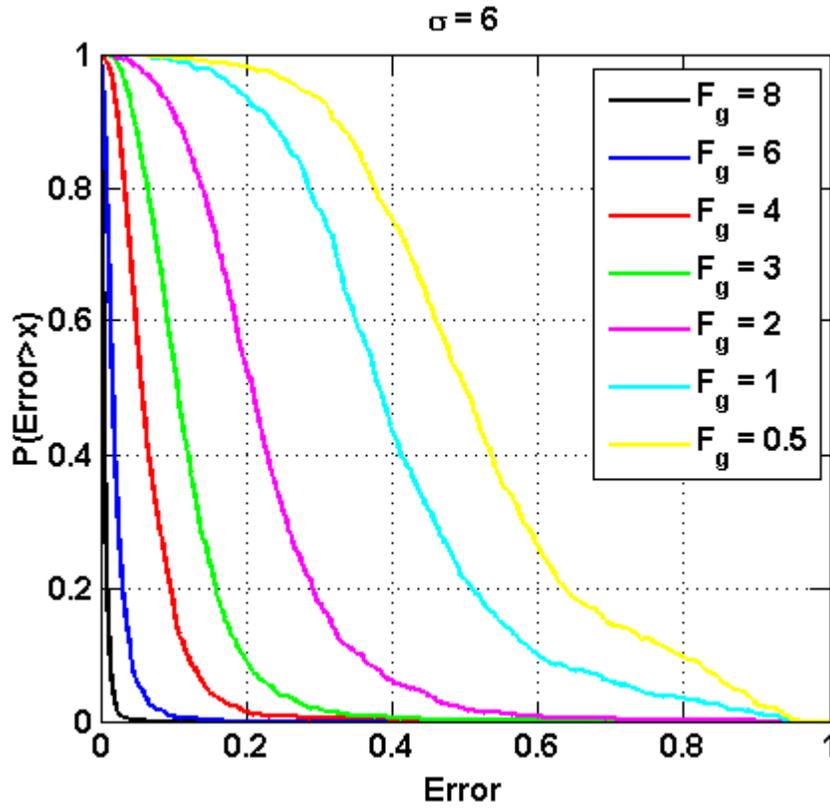


Figure 3.4: Cumulative distribution function of the convergence error E_t given a fixed number of algorithm iterations determined by different values of the gossip iteration factor F_g .

cumulative distribution function (CDF) of E_t and determining the largest value of E_t that had a cumulative probability smaller than itself. Figure 3.4 shows a plot of the CDF of E_t for several values of F_g . The figure shows that the error E_t approaches zero with high probability as the number of iterations increases, a behavior that supports the results in Figures 3.2 and 3.3.

3.5 Energy Consumption Analysis

In this section the energy consumption of the MTWARL distributed localization algorithm is compared to that of the centralized localization algorithms presented in previous sections. The average transmit energy is used as the measure of energy consumption. This metric

is often used to quantify energy consumption in WSN/CR applications, where the nodes are portable battery-operated devices and most of the energy is consumed by the transmitter. Specifically, the average transmit energy required for producing one set of transmitter location estimates is considered. This metric is given by

$$\begin{aligned} E &= \frac{P_d^T B K}{R} \\ &= P_d^T B K, \end{aligned} \tag{3.25}$$

where P_d^T is the power transmitted such that the minimum received power for link closure is received at distance d from the transmitter, R is the bit rate, K is the number of values that are transmitted between nodes, and B is the number of bits used to represent each of the values to be transmitted. As the equation shows, for simplification, the bit rate R is assumed to be equal to 1 for all transmissions. Of the three variables in the right-hand side of equation (3.25), only K and P_d^T are dependent on the type of algorithm. Therefore, they are explored and expanded to derive expressions of E for the algorithms under evaluation. B is treated as a design parameter and is carried out through the analysis. In what remains of this section, expressions for E will be derived for the centralized and distributed multiple transmitter localization algorithms.

3.5.1 Energy Consumption of Centralized MTL Algorithms

In this section, the average transmit energy of a centralized multiple transmitter localization algorithm is derived. Even though two of such algorithms have been discussed, no distinction is made in computing their energy utilization because of their identical transmission schemes. We begin the analysis with K , to which a subscript c is added to denote its reference to a centralized localization algorithm. In RSS-based centralized localization algorithms, each node sends three values to the fusion center: the received power measurement $r_{j(dB)}$, and the node's Cartesian coordinates (x, y) . The fusion center sends back a

pair of Cartesian coordinates for each transmitter back to the nodes, for a total of $2M$ values. Therefore, the total number of values transmitted for localizing and returning the results back to the nodes is

$$K_c = (3 + 2M)N. \quad (3.26)$$

Next, the the value of P_d^T is determined. For the purpose of this analysis, a simplified expression of the relationship between received and transmit power is utilized. This removes factors such as frequency of operation and antenna characteristics, which are similar for most systems and are not as influential to the transmit power as distance is. Such relation is given by

$$P_d^T = \frac{P_{min}d^\gamma}{\rho} \approx d^\gamma, \quad (3.27)$$

where P_{min} is the minimum received power to establish reliable communication between nodes. As (3.27) shows, P_d^T is proportional to d^γ .

Since we are interested in average energy consumption, we compute P_d^T relative to the average distance between nodes \bar{d} . We employ probabilistic analysis to compute \bar{d} . Consider a squared area of interest with sides of length L . Consider two nodes ϕ_1 and ϕ_2 that have coordinates (x_1, y_1) and (x_2, y_2) , respectively. Assume these coordinates are realizations of a random variable uniformly distributed in the range $(0, L)$. The distance between these two nodes is given by

$$d(\phi_1, \phi_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}. \quad (3.28)$$

Since $d(\phi_1, \phi_2)$ is a function of random variables (x_1, y_1) and (x_2, y_2) , it is also a random variable. Employing geometric reconstruction several times for each of the operations in equation (3.28), we can compute the cumulative distribution function (CDF) of $d(\phi_1, \phi_2)$

as

$$F_d(d) = \begin{cases} 0, & d < 0 \\ \frac{\pi}{L^2}d^2 - \frac{8}{3L^3}d^3 + \frac{1}{2L^4}d^4, & 0 \leq d < L \\ 1 - \left[\frac{2}{3} + 2b + \frac{b^2}{2} - \frac{2}{3}\sqrt{(b-1)^3} \right. \\ \quad \left. - 2\sqrt{b-1} - 2b\sqrt{b-1} - 2b \arcsin \frac{2-b}{b} \right], & L \leq d < L\sqrt{2} \\ 1, & d \geq L\sqrt{2} \end{cases} \quad (3.29)$$

where $b = d^2/L^2$. Taking the derivative of (3.29) to obtain the PDF, we obtain

$$f_d(d) = \begin{cases} 0, & d < 0 \\ \frac{2d}{L^4}(\pi L^2 - 4Ld + d^2), & 0 \leq d < L \\ \frac{d}{L^6 - L^4 d^2} \left[L^2 d^2 \left(-6\sqrt{b-1} + 4\sqrt{\frac{(d^2-L^2)}{bd^2}} + 2 \right) \right. \\ \quad \left. + (4L^2 d^2 - 4L^4) \arcsin(1 - 2/b) \right. \\ \quad \left. + L^4 \left(2\sqrt{b-1} - 2\sqrt{\frac{(d^2-L^3)^3}{L^6}} - 4 \right) + 2d^4 \right], & L \leq d < L\sqrt{2} \\ 1, & d \geq L\sqrt{2} \end{cases} \quad (3.30)$$

Figure 3.5 shows a comparison between the PDF of $d(\phi_1, \phi_2)$ obtained by evaluating equation (3.30) and that obtained by a Monte Carlo simulation. Given the $f_d(d)$, the expected distance between nodes can be determined by evaluating the expectation integral $\bar{d} = \int l f_l(l) dl$, where l is a dummy integration variable. Unfortunately, the term over $L \leq d < L\sqrt{2}$ in (3.30) is not integrable. To circumvent this problem, the term over

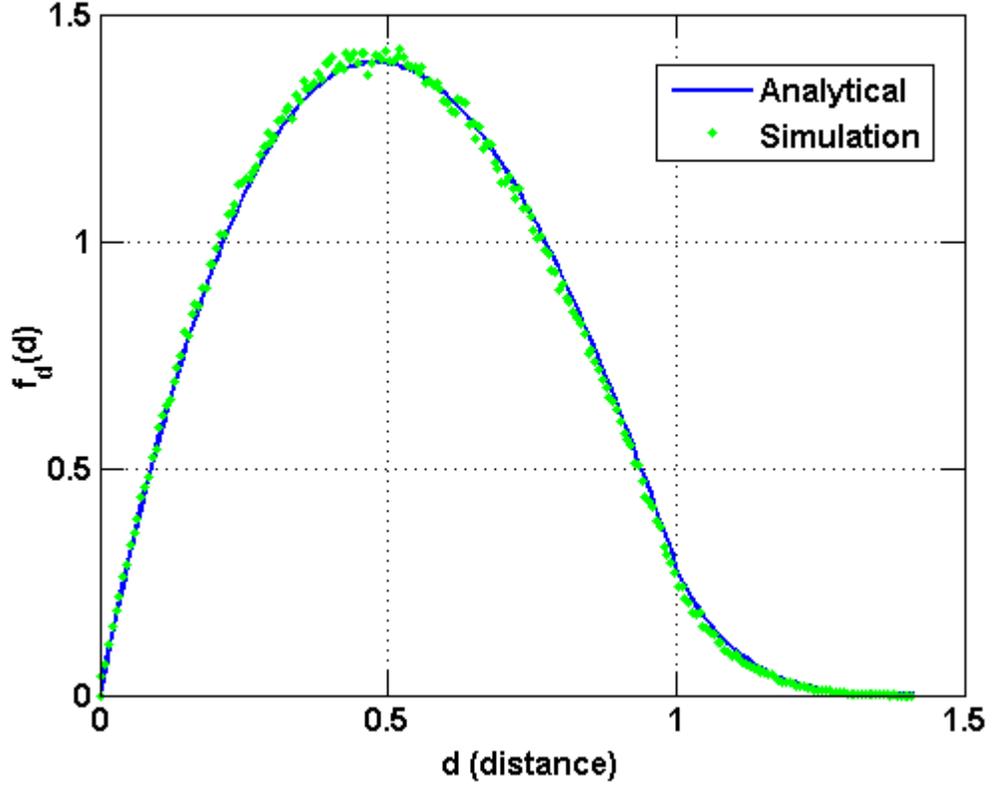


Figure 3.5: Comparison of analytical and simulation PDF of $d(\phi_1, \phi_2)$ in a unit square area ($L = 1$).

$L \leq d < L\sqrt{2}$ is approximated by half the area of the expectation integral of a line connecting points $(L, f_d(L))$ and $(\sqrt{L}, 0)$. Given this approximation, \bar{d} is computed as

$$\begin{aligned}
 \bar{d} &= \int_0^\infty l f_l(l) dl \\
 &\approx \int_0^L \frac{2l^2}{L^4} (\pi L^2 - 4Ll + l^2) dl + \frac{f_d(L)}{2L(\sqrt{2}-1)} \int_{\sqrt{2}L}^L \sqrt{2}Ll + -l^2 dl \\
 &= 0.4944L + 0.1179f_d(L)L^2.
 \end{aligned} \tag{3.31}$$

Numerical simulations reveal that the approximation error of (3.31) is within 1% of the

actual solution. Given \bar{d} and K_c , the average transmit energy can be determined as

$$\begin{aligned} E_c &= P_{d_c}^T B K_c \\ &= (0.4944L + 0.1179f_d(L)L^2)^\gamma B(3 + 2M)N. \end{aligned} \tag{3.32}$$

3.5.2 Energy Consumption of the MTWARL Distributed MTL Algorithm

In this section, the average transmit energy of the MTWARL algorithm is considered. As detailed in Section 3.3, the MTWARL algorithm consists of a outer (or main) loop and a inner (or secondary) loop. During each iteration of the main loop, clustering is performed based on proximity to the latest location estimate at each cluster. In addition, inner iterations perform distributed localization across the nodes in each newly formed cluster. Since each clusters localizes a single transmitter, once the main loop iterations are completed, the results are disseminated across clusters. Given this information, we define the total average transmit energy for the distributed algorithm as

$$E_g = \sum_{k=1}^I \left(P_{d(k)}^T B K_{g(k)} \right) + E_{ret}, \tag{3.33}$$

where $P_{d(k)}^T$ and $K_{g(k)}$ are the required minimum transmit power for link closure and number of transmitted values for the k th main loop iteration of the algorithm, respectively, and I is the total number of iterations (outer + inner loop iterations). E_{ret} is the energy required to disseminate the results across clusters.

We begin by computing the average required transmit power $P_{d(k)}^T$ for each iteration. For each main loop iteration, the algorithm uses the receivers contained inside a circular area centered around the latest transmitter estimate. The computation of the average distance between receivers within a circular area requires a coordinate conversion and the

evaluation of several complicated integrals to compute the density function and to evaluate the expectation integral. Since the shape of the distribution is not of interest, we resort to a simple approximation based on equation (3.31). This method approximates the average distance in a circular region by the average distance in a squared region with sides equal to $L_k \left(\frac{1+\sqrt{2}}{2\sqrt{2}} \right)$. This area corresponds to the square who is inscribed at the half way point between the squares with sides equal to L_k and $\frac{L_k}{\sqrt{2}}$. Given this information, the average distance between nodes at the k th iteration is approximated as

$$\bar{d}_k \approx 0.4944L_k \left(\frac{1+\sqrt{2}}{2\sqrt{2}} \right) + 0.1179f_d \left(L_k \left(\frac{1+\sqrt{2}}{2\sqrt{2}} \right) \right) L_k^2 \left(\frac{1+\sqrt{2}}{2\sqrt{2}} \right)^2, \quad (3.34)$$

$$= 0.4220R_k + 0.0859f_d(0.8536R_k) R_k^2, \quad (3.35)$$

where $R_k = L_k/2$ is the clustering radius of the k th iteration and $f_d(0.8536R_k)$ is the density defined in (3.30) with parameter $L = 0.8536R_k$ and evaluated at $d = 0.8536R_k$. Substituting (3.34) into (3.27) to determine $P_{d(k)}^T$ we obtain

$$P_{d(k)}^T = (0.4220R_k + 0.0859f_d(0.8536R_k) R_k^2)^\gamma. \quad (3.36)$$

Similarly to $P_{d(k)}^T$, the number of values that are transmitted also varies with the iteration number and is defined as $K_{g(k)}$. In the MTWARL algorithm, each message transmitted between nodes carries a total of 6 values: 1 two-dimensional value (2 real values) for the numerator of (3.8), 1 value for denominator of (3.8), 1 value for the gossip iteration counter, and 2 values to indicate the location of the center of the cluster. Given this information, the number of transmitted values per inner loop iteration is given by

$$K_{g(k)} = 6I_g = 18N_k, \quad (3.37)$$

where $I_g = 3N_k$ is the number of gossip iterations required for convergence without significant accuracy loss as determined in Section 3.4 and N_k is the average number of receivers per cluster at the k th iteration. Since the receivers are placed according to a uniform distribution across the area of interest, the value of N_k is determined by the portion of the area of interest that the cluster area takes. This relation is expressed as

$$N_k = \text{round} \left(\frac{\pi A_k N}{4A} \right) \quad (3.38)$$

$$= \text{round} \left(\frac{\pi N R_k^2}{A} \right), \quad (3.39)$$

where A_k is the area of a square with side length $2R_k$ and A is the area of the region of interest. Table 3.1 shows typical values of cluster radius, average distance between nodes and expected number of receivers per cluster N_k for each iteration in the case where $M = 2$.

Table 3.1: Two transmitter energy consumption parameters per iteration.

Global Iteration	Tx being Localized	R_k (meters)	\bar{d}_k (meters)	N_k
1	1	0.4	0.3615	$\text{round}(0.5N)$
2	1	0.3	0.2711	$\text{round}(0.27N)$
3	1	0.3	0.2711	$\text{round}(0.27N)$
4	2	0.4	0.3615	$\text{round}(0.5N)$
5	2	0.3	0.2711	$\text{round}(0.27N)$
6	2	0.3	0.2711	$\text{round}(0.27N)$

Next, the energy required for dissemination of the estimates is considered. Since the members of each cluster are aware of who the other members of their cluster are, edge receivers send their results to the nearest non-member of their cluster. This node then starts a gossip algorithm instance to disseminate the results in its cluster. Given this

dissemination scheme, the dissemination energy E_{ret} is expressed as

$$\begin{aligned} E_{ret} &= 2B\bar{d}_{min}^{\gamma}\text{round}((1 - (1/M))N) \\ &+ 12MB(0.3L)^{\gamma}\text{round}(\pi(0.3L)^2N/100), \end{aligned} \quad (3.40)$$

where \bar{d}_{min} is the average minimum distance between receivers. Equation 3.40 shows that the energy for dissemination has two additive components: 1)the communication between the edges of the clusters, and 2)the actual dissemination over the cluster. The edge communication is computed as the energy required for transmitting two result values across the average number of receivers that adjoin other clusters. The dissemination component represents the energy required to run a gossip algorithm instance within each cluster. Substituting results of equations (3.36), (3.37), (3.40) and (3.38) into (3.33) we obtain

$$\begin{aligned} E_g &= 2B\bar{d}_{min}^{\gamma}\text{round}((1 - (1/M))N) \\ &+ 12MB(0.3L)^{\gamma}\text{round}(\pi(0.3L)^2N/100) \\ &+ 18B \sum_{k=1}^I (0.4220R_k + 0.0859f_d(0.8536R_k)R_k^2)^{\gamma} \text{round}\left(\frac{\pi NR_k^2}{A}\right). \end{aligned} \quad (3.41)$$

3.5.3 Energy Utilization Comparison Between Centralized and Distributed Approaches

Figure 3.6 shows a comparison between the centralized and distributed processing algorithms considered in the previous sections. The normalized average transmit energy is plotted versus N for $M = 2$ and $\gamma = 2$. Curves for two versions of the MTWARL distributed algorithm are shown, where the difference between the two lie in the way the initial clustering is performed during the first main loop iteration ($i = k = 1$). Version 1 consists of the simplest approach where all the nodes are included in the initial cluster. On the other hand, version 2 consists of the more efficient approach, where the initial cluster is

performed over a small region centered on the center of the area of interest. As the figure shows, average transmit energy consumption increases linearly with N . The curves for the distributed algorithms exhibit discontinuities due to the round operation that is part of the computations. The figure also shows that for this value of γ , the centralized localization

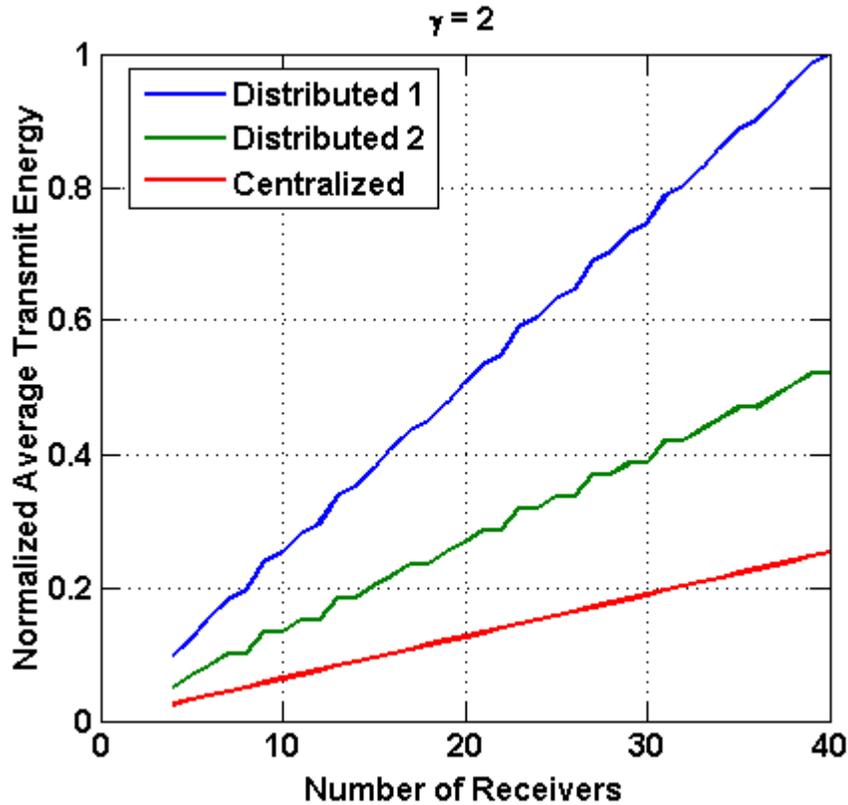


Figure 3.6: Normalized average transmit energy of two versions of the MTWARL distributed localization algorithm compared to their centralized counterpart for $M = 2$ and $\gamma = 2$.

algorithm consumes the least amount of energy and version 2 of the distributed algorithm consumes approximately 20% more energy than the centralized algorithm. Version 1 of the distributed algorithm consumes the most power since its first iteration gossips over all N receivers across the entire area, sending a large number of messages over the longest possible distance. These results underscore how clustering and local gossiping (within clusters) are critical to lowering the total energy consumption of the system.

As equations (3.32) and (3.41) show, the average transmit energy varies as a function of

γ . Figure 3.7 shows the average transmit energy of version 2 of the distributed algorithm relative to that of the centralized algorithm for $\gamma = 2, 3, 4$ and 6 . The figure shows that the relative energy consumption of the distributed algorithm varies dramatically with γ , and that it ranges from approximately 105% higher energy consumption at $\gamma = 2$, to approximately 45% of the energy consumption of the centralized algorithm at $\gamma = 6$. These results

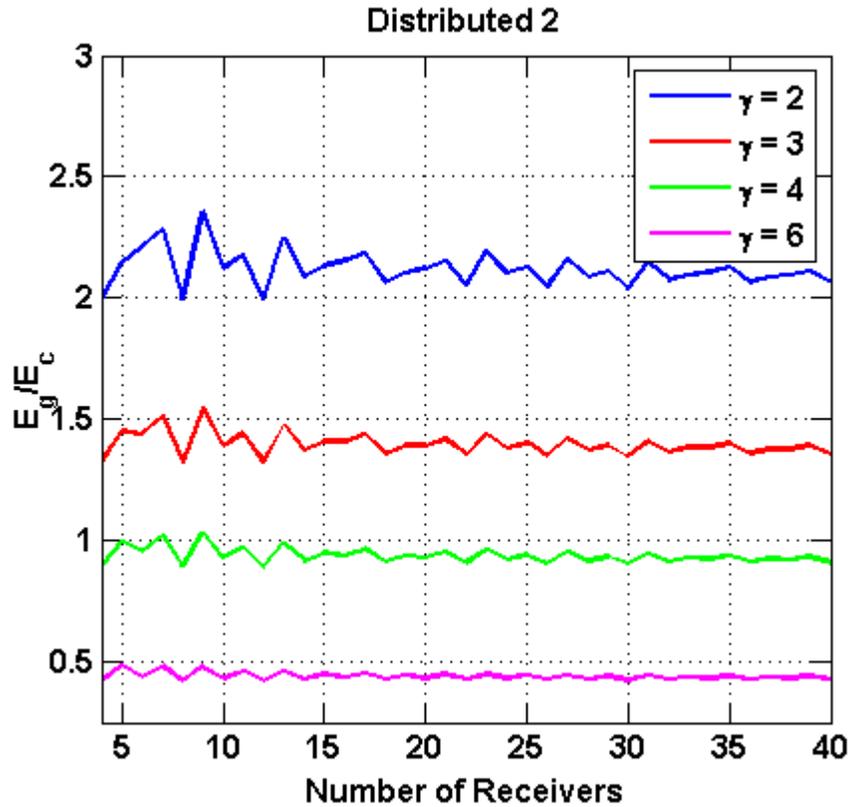


Figure 3.7: Average transmit energy of the distributed algorithm version 2 relative to the energy consumption of the centralized algorithm for $\gamma = 2, 3, 4$, and 6 .

suggests that if energy is constrained, the use of randomized distributed processing algorithms would only be advantageous in highly obstructed propagation environments where γ is relatively high. However, if energy is not constrained, the fault tolerance provided by distributed processing would be traded for an increase of up to 105% in energy consumption. The results also underscore that eventhough the considered distributed algorithms transmit over shorter distances, their total energy consumption may be larger since more

transmissions of more data values are performed. This is an inherent trade of distributed algorithm design that is worsen by the randomized nature of the gossip algorithm, which avoids complicated routing schemes at the expense of longer convergence times, resulting in larger amounts of transmitted information.

3.6 Computational Complexity Analysis

In this section, the computational complexity of the MTWARL distributed algorithm is compared to that of the centralized LCIS algorithm. The objective of this comparison is to assess the suitability of the MTWARL algorithm to be deployed in an WSN/CR system, where the nodes have limited computational capabilities and computational complexity is critical. The LCIS was chosen for the comparison as it is the ML-based centralized algorithm with the lowest complexity, thus offering an appropriate reference point to draw comparisons from. The metric for computational complexity used is the floating point operations (FLOP). This metric is typically used in the field of scientific computing for measuring the computational complexity of a given calculation.

3.6.1 Computational Complexity of Centralized MTL Algorithms

As discussed in previous sections, the majority of the computational complexity of the ML-based algorithms is contained in the non-linear minimizations used to obtain the different location estimates. Recall the number of minimizations that are performed within the LCIS algorithm is given by

$$O_{IS} = 2M. \tag{3.42}$$

The Nelder-Mead (NM) direct search method is used for performing these minimizations [53]. As in any direct search method, the NM method only uses values of the cost function to search for a solution and thus the computational complexity is mostly driven by the complexity of the cost function, which is evaluated twice per iteration [52]. Recall the cost

function for the LCIS is given by

$$\hat{\theta}_i = \arg \min_{\theta_i} \sum_{j=1}^K (\hat{r}_{ij(dB)} - E[r_{ij(dB)}|\theta_i])^2, \quad (3.43)$$

where $E[r_{ij(dB)}|\theta_i] = P_i^T \rho \left(\frac{d_0}{d_{ij}(\theta_i)} \right)^\gamma$. Given this information, the number of FLOPs per cost function evaluation can be computed by adding the $9N$ FLOPs required to compute $E[r_{ij(dB)}|\theta_i]$ to the $3N$ FLOPs required to compute summation in (3.43). Thus, the computational complexity of evaluating the cost function is given by

$$C_{eval} = 12N. \quad (3.44)$$

Given this information, the complexity of a NM minimization is approximated by

$$\begin{aligned} C_{nm} &\approx 2I_{nm}C_{eval} \\ &= 24I_{nm}N, \end{aligned} \quad (3.45)$$

where I_{nm} is the number of NM iterations required for convergence. Having obtained the number of minimizations from (3.42) and the complexity of each minimization from (3.45), the complexity of computing an estimate via the LCIS algorithm can be expressed as

$$\begin{aligned} C_{cen} &= O_{IS}C_{nm} \\ &= 48I_{nm}NM. \end{aligned} \quad (3.46)$$

Empirical results have revealed that typically $I_{nm} \approx 150$.

3.6.2 Computational Complexity of the MTWARL Distributed MTL Algorithm

In this section, the computational complexity of the MTWARL algorithm is considered. The complexity analysis follows the formulation established in Section 3.5.2 for the analysis of the algorithm's energy consumption. The MTWARL algorithm has two main computations: 1) the computation of the distances performed as part of the clustering step within each of the main loop iterations, and 2) the transmitter location update that is performed between pairs of receivers during every gossip iteration. Given this information, the complexity of the MTWARL algorithm can be computed as

$$C_{dis} = C_1 + C_2, \quad (3.47)$$

where C_1 and C_2 are the complexities of the distance calculations and the transmitter location updates, respectively, as enumerated above.

The distance computations performed during each iteration are of the form

$$d(\theta_j, \phi_i) = \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2}, \quad (3.48)$$

and, as the equation shows, requires 6 FLOPs for its computation. During each of the I main loop iterations, this computation is performed once per receiver available for clustering. Therefore, according to the algorithm, the C_1 component can be expressed as

$$\begin{aligned} C_1 &= 6N + 6 \sum_{k=1}^{I-1} N_k \\ &= 6N + 6 \sum_{k=1}^{I-1} \text{round} \left(\frac{\pi N R_k^2}{A} \right), \end{aligned} \quad (3.49)$$

where N_k is the average number of receiver used for estimation during the k th main loop

iteration.

The transmitter location update is performed as a part of the pairwise gossip algorithm employed by the MTWARL algorithm. This update is performed by separately updating the numerator and denominator of the estimate expression defined in equation (ref to Ch.3). On the v th gossip iteration, the p th and q th receivers compute their update according to the following two equations:

$$\hat{\theta}_{num}(v) = \frac{\phi_p r_p + \phi_q r_q}{2} \quad \hat{\theta}_{den}(v) = \frac{r_p + r_q}{2}, \quad (3.50)$$

where $\hat{\theta}_{num}(v)$ is the numerator portion and $\hat{\theta}_{den}(v)$ is the denominator portion. As the equation shows, the computational complexity of a single update computation (both numerator and denominator) is 8 FLOPs. Since both of the gossiping receivers update their location estimate, equation (3.50) is computed twice on every gossip iteration. Additionally, an instance of the gossip algorithm is spawned during each of the I iterations of the main loop. Given this information, the component C_2 can be expressed as

$$\begin{aligned} C_2 &= \sum_{k=1}^I 16I_g \\ &= 48 \sum_{k=1}^I \text{round} \left(\frac{\pi N R_k^2}{A} \right), \end{aligned} \quad (3.51)$$

where $I_g = 3N_k$ is the typical number of gossip iterations required for convergence. Substituting (3.49) and (3.51) into (3.47), the expression for the computational complexity of the MTWARL algorithm can be expanded as follows:

$$\begin{aligned} C_{dis} &= C_1 + C_2 \\ &= 6N + 6 \sum_{k=1}^{I-1} \text{round} \left(\frac{\pi N R_k^2}{A} \right) + 48 \sum_{k=1}^I \text{round} \left(\frac{\pi N R_k^2}{A} \right), \end{aligned} \quad (3.52)$$

3.6.3 Computational Complexity Comparison Between Centralized and Distributed Approaches

Figure 3.8 shows a computational complexity comparison between the centralized and distributed multiple transmitter localization algorithms under consideration. The figure shows the analytical expressions (3.46) and (3.52) as well as empirical benchmarks of the algorithms versus the total number of receivers N . In both the analytical and empirical cases, the results were normalized with respect to the maximum complexity. This allows a comparison of relative algorithm computational complexity between analytical and empirical results. The empirical benchmarks were run in Matlab and consisted of measuring the elapsed time of 500 iterations of each algorithm. Both algorithms were run sequentially within the same script to ensure equal runtime environments. As the plot shows, the computational complexity of the distributed algorithm is much smaller than that of the centralized algorithm, a trait that comes at the expense of lower localization accuracy. The results confirm that the proposed distributed algorithm is well suited for a distributed implementation where the nodes are likely to have limited computational capabilities. The plot also shows that the empirical and analytical results are in agreement with one another, corroborating the validity of the analysis. It is worth noting that the empirical and analytical curves for the centralized algorithm differ by approximately 17% at $N = 10$, and converge to approximately the same value at $N = 40$. This is due to the termination test of the NM search algorithm (which is constant in the number of receivers) starts to be insignificant as the number of receiver increases, and the cost function evaluations dominate the total computational complexity of the algorithm.

3.7 Localization Accuracy Comparison

In this section, the localization performance of the MTWARL distributed algorithm is compared to that of the LCIS and IS-based EM centralized algorithms. Thus far, a trade off

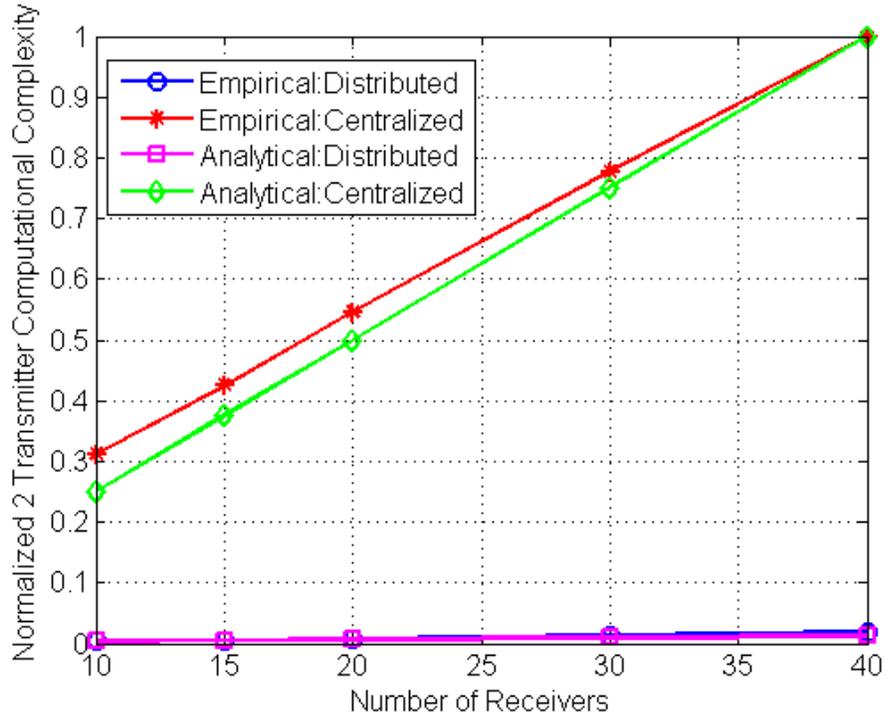


Figure 3.8: Two-transmitter-localization normalized computational complexity comparison between centralized and distributed localization algorithms.

between computational complexity and localization accuracy has been studied for centralized multiple transmitter localization algorithms. The objective of the comparison in this section is to extend the trade space to include properties such as energy utilization and fault tolerance that are introduced by distributed processing, and explore any trade offs they may have with localization accuracy. The comparison is performed via numerical simulations, in which the results obtained in Section 3.4.2 are used to configure the proposed distributed algorithm for energy efficient operation.

Simulations for all the algorithms follow the same general setup described in Section 2.6. The IS-based EM algorithm was run with $C = 2$ and $I = 8$. This represents 4 times the computational complexity of the LCIS and 320 times the computational complexity of the MTWARL. The MTWARL was setup according to table 3.1 with gossip stochastic matrix set to P^* , as defined in (3.18). The number of gossip iterations was set to a fixed

value according to $I_g = 3N_i$. This relation was obtained from the results of Section 3.4.2, where the impact of I_g to localization accuracy was studied. Figure 3.9 shows the simulated performance for the algorithms. As with previous localization performance simulations, the chosen performance metric is the average squared distance error between estimated and true transmitter locations, where the average is taken over the M transmitter location estimates. Performance figures show both the median and mean error over 1000 different random draws of $M = 2$ and 3 transmitters, and $N = 4$ to $N = 40$ receivers with shadowing variance of $\sigma^2 = 16, 36$ and 64.

As the figures show, the IS-based EM produces the smallest mean and median errors across all values of N and σ^2 considered, while the MTWARK produces the largest mean and median errors. The performance gap between the MTWARK and the other two centralized algorithms widens with increasing σ^2 values, suggesting that the MTWARK is more susceptible to perturbations caused by the random nature of lognormal shadowing. This gap is also seen to decrease with N , specially for $\sigma^2 = 16$ and 36, indicating that that MTWARK is not well suited for systems with few sensor nodes. This is believed to be due to the MTWARK not being able to refine its estimate because of a lack of receivers in the clustering area around a location estimate.

The comparison of localization accuracy reveals that computational complexity, energy consumption and fault tolerance could be traded for localization performance. Table 3.2 shows a summary of the properties of the algorithms under consideration. From the comparison it is evident that the low complexity, low energy consumption and fault tolerance of the proposed MTWARK distributed algorithm come at the cost of lesser localization accuracy, specially on environments with severe shadowing and in systems with low number of sensors. On the other hand, high localization accuracy can be achieved by significantly increasing the computational complexity and moderately increasing the energy consumption.

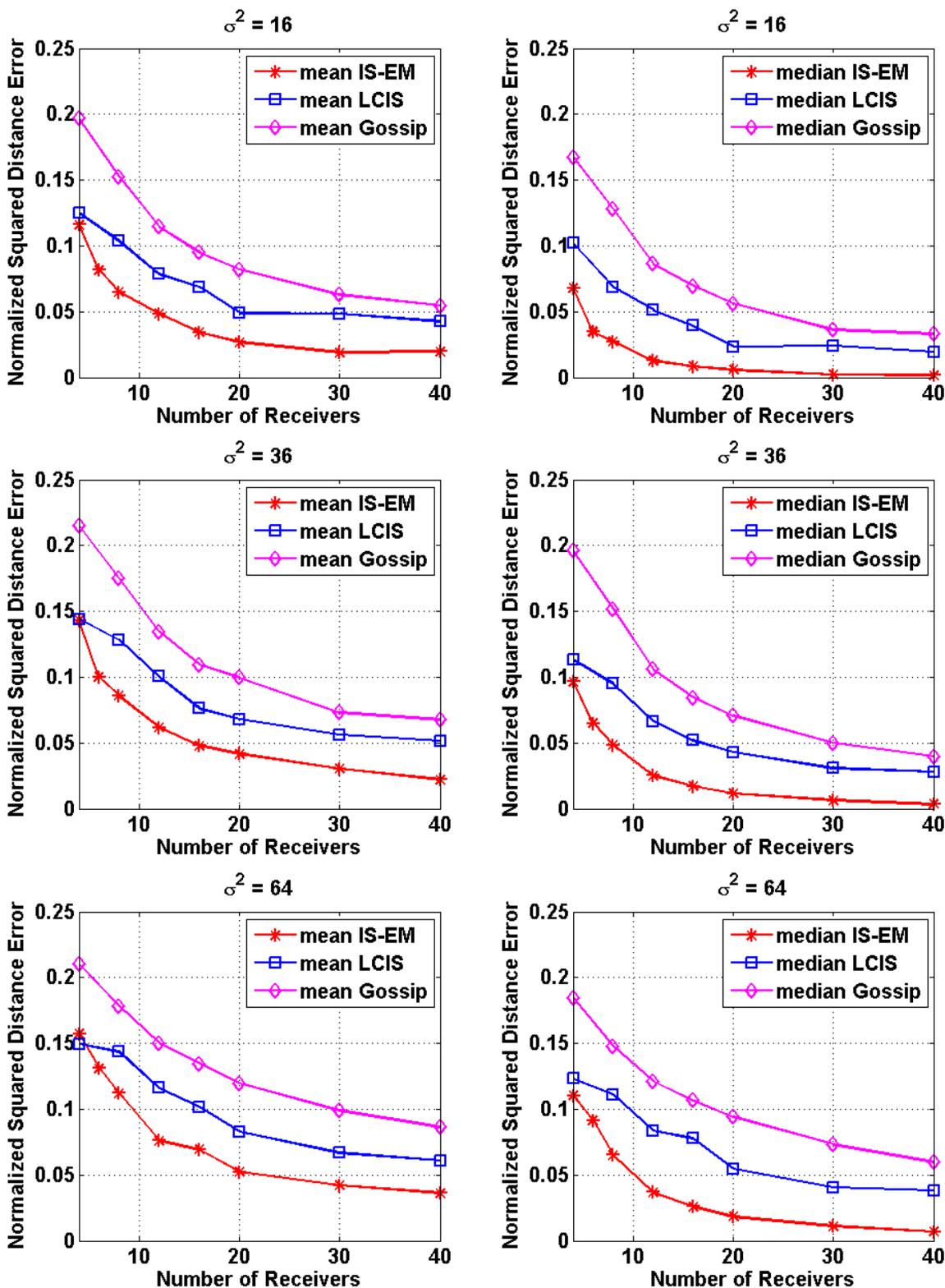


Figure 3.9: LCIS, IS-EM and MTWARL two transmitter mean (left column) and median (right column) localization accuracy for $\sigma^2 = 16$ (top row), 36 (middle row) and 64 (bottom row).

Table 3.2: Comparison of multiple transmitter localization algorithms.

Algorithm	Fault Tolerance	Energy Consumption	Computational Complexity	Localization Accuracy
ISEM	No	Med	Very high	High
LCIS	No	Med	High	Med
MTWARL	Yes	Low-to-Med	Low	Low

Chapter 4: Effects and Exploitation of Sensor Mobility in MTL

One major characteristic of wireless mobile networks is user mobility. Systems such as CRNs are comprised of users that may move about an area of operation. This mobility affords a sensor the ability to take measurements at different locations. These additional measurements contain information that can be used to complement previous measurements, and thus improve localization accuracy. Thus far, we have studied MTL assuming that sensors are stationary at fixed locations. In this chapter, we ease the constraints on receiver mobility and explore its effects on the statistical properties of RSS observations and its implications on ML-based MTL. In addition, we investigate ways of taking advantage of the additional information contained in measurements taken by receivers after moving from their original locations.

4.1 Extensions to System Model

4.1.1 Mobility Model

Mobility is a topic that has received significant attention in the wireless network literature, especially in the areas of mobile cellular and mobile ad hoc networks [65, 66]. Several mobility models have been studied in order to describe movement patterns of mobile users, and how their location, velocity and acceleration vary as a function of time. In mobile cellular networks, for example, the models focus on the mobility of users relative to an area of operation (e.g., a cell), which allows the study of coverage, cell change rates, and hand overs. In contrast, in mobile ad hoc networks, the models focus on user movement relative to other users, as this facilitates the study of wireless connectivity and optimal routing protocols. Since CRN's are essentially a type of ad hoc network, the focus of

mobility models for both types of networks significantly overlap. We draw from previous work in mobility models for mobile ad hoc networks and employ the random waypoint model (RWM) as a means of modeling movement [67]. The RWM falls within the realm of random mobility models. In these models, nodes move independently according to a random movement scheme (e.g., random destination, random direction, random walk, etc.). In RWM, each node moves independently to a random destination with a random constant velocity. Due to its simplicity, the RWM has been widely used in mobile ad hoc networks research. Below we expand on the inner workings of the RWM.

Let $\mathcal{S} = \{j\}_{j=1}^N$ be the set containing the indices of all N sensor nodes performing MTL. The initial location of the N sensor nodes is denoted by the matrix $\Phi = [\phi_1; \phi_2; \dots; \phi_N]$, where ϕ_j is the location of the j th receiver, which is given by the Cartesian coordinate pair (x_j, y_j) . Under the RWM, receivers independently choose their new locations, which are given by

$$\Phi^{(n)} = [\phi_1^{(n)}; \phi_2^{(n)}; \dots; \phi_N^{(n)}], \quad (4.1)$$

where n is an integer that denotes the n th element in the sequence of sensor location matrices $\Phi^{(n)}$. Each receiver travels to its respective destination with a random constant velocity drawn from a uniform distribution over the interval $[V_{min}, V_{max}]$, where V_{min} and V_{max} are the minimum and maximum velocities with which a receiver will travel, respectively. The range of motion is controlled by the probability density function of the destination, which is typically chosen as two-dimensional uniform over the intervals $[0, x_{max}]$ and $[0, y_{max}]$. Since receivers are assumed stationary at the time a measurement is taken, velocity is assumed high enough such that each sensor reaches its destination before a new measurement is taken. This assumption results in the following minimum velocity:

$$V_{min} \geq d_{max}/T_m, \quad (4.2)$$

where $d_{max} = \sqrt{x_{max}^2 + y_{max}^2}$ and T_m is the measurement interval. Once receiver nodes

reach their destinations, they all take new measurements. The new set of measurements is given by

$$\mathbf{r}_{(dB)}^{(n)} = [r_{1(dB)}^{(n)} r_{2(dB)}^{(n)} \cdots r_{N(dB)}^{(n)}]^T, \quad (4.3)$$

where $r_{j(dB)}^{(n)}$ denotes the n th observation of received power at the j th sensor node measured in dB. In subsequent sections we will discuss how these new measurements can be used to improve previously computed location estimates.

It is worth mentioning that analysis of the asymptotic behavior of the RWM with time has revealed that as time elapses, the node spatial distribution across the area of operation oscillates between a uniform distribution to one that is highly concentrated near the center of the area [68]. This phenomenon would negatively affect localization performance if RWM were used to model long-term mobility behavior. However, since the interest in using RWM is to model mobility over short periods of time (a maximum of one to two waypoint sets), this phenomenon does not manifest itself in our models and thus does not corrupt localization accuracy.

4.1.2 Correlated Shadowing Model

One important effect to take in to account when studying mobility is the correlation among received power measurements taken at relatively short distances from each other. This correlation arises from the fact that signals undergo similar propagation effects (i.e, blockages, reflections, and scattering) when the measuring receivers are located near each other. In our study, we assume an empirically derived model known in the literature as the Gudmundson model [40, 69]. In this model, shadowing is modeled as a first-order autoregressive process, which implies an autocorrelation function that decays exponentially with distance between the measurements. This results in covariance matrix Σ , whose elements are given by

$$\{\Sigma\}_{pq} = \sigma^2 \rho_{pq} = \sigma^2 \exp\left(\frac{-d_{pq}}{D_c}\right), \quad (4.4)$$

where ρ_{pq} is the correlation coefficient relating the measurement from the p th and the q th receiver nodes, d_{pq} is the distance between the p th and the q th receiver nodes, and D_c is the correlation distance. The correlation distance can be defined as the distance at which the correlation coefficient equals 0.3679 and is determined based on the propagation environment. A typical value of D_c for a suburban environment is 500 meters, whereas a value of 20 meters is typical for an urban environment [40, 70]. The Gudmundson model is widely accepted in mobile wireless research, and is the basis for other models that have been proposed [69–71].

4.2 Implications of Correlated Shadowing to ML-based MTL Framework

As discussed in Chapter 2, the derivation of the ML localization estimate assumes that the lognormal shadowing affecting each transmitter-receiver link is independent. This assumption can be considered valid as long as there is enough separation among receivers [40]. When considering sensor mobility, it is of interest to consider scenarios in which sensors move in an uncontrolled and random manner (e.g., a sensor mounted on a moving vehicle, or on a mobile device that a pedestrian is carrying). In these scenarios, sensors may move to locations that violate the proximity assumption made for assuming independent lognormal shadowing and thus, may take measurements that are correlated to the ones taken at nearby locations. In this section, we explore the implications of such scenarios to the ML estimate formulation discussed in Section 2.3.1.

Recall the likelihood function of the vector of RSS observations $r_{(dB)}$ containing power from a single transmitter, given in equation (2.20) as

$$f(r_{(dB)}|\theta) = \prod_{j=1}^N \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(\frac{-(r_{j(dB)} - \mu_j(\theta))^2}{2\sigma_j^2}\right).$$

Because of the independence assumption, (2.20) is comprised of the product of the density functions of the individual measurements $r_{j(dB)}$. When the measurements in $r_{(dB)}$ are dependent, due to the correlated shadowing, the likelihood function is a multivariate normal distribution given by,

$$f(r_{(dB)}|\theta) = \frac{1}{(2\pi)^{N/2} |\Sigma|^{1/2}} \exp - \left(\frac{1}{2} (r_{(dB)} - \mu(\theta))^T \Sigma^{-1} (r_{(dB)} - \mu(\theta)) \right), \quad (4.5)$$

where $\mu(\theta) = [\mu_1(\theta), \mu_2(\theta), \dots, \mu_N(\theta)]$ contains the expected received power at each sensor. Equation (4.5) can be manipulated akin to equation (2.20) to obtain the log-likelihood function. Substituting the log-likelihood function into equation (2.18), the ML estimate is obtained by solving

$$\hat{\theta}_{ML} = \arg \min_{\theta} \left((r_{(dB)} - \mu(\theta))^T \Sigma^{-1} (r_{(dB)} - \mu(\theta)) \right). \quad (4.6)$$

As equation (4.6) shows, the inverse of covariance matrix of the shadowing Σ smears each receiver's contribution to account for the correlation among measurements. Another important observation about (4.6) is that the cost function to be minimized is significantly more computationally complex than the cost function of (2.21), requiring an additional matrix-vector multiplication ($[1 \times N][N \times N]$ or $[N \times N][N \times 1]$). These additional computations are equivalent to N^2 additional FLOPS per cost function evaluation, which represent an increase in complexity given by

$$\begin{aligned} \Delta_C &= \frac{C_{eval}^s}{C_{eval}} - 1 \\ &= \frac{12N + N^2}{12N} - 1 \\ &= \frac{N}{12}, \end{aligned} \quad (4.7)$$

where $C_{eval} = 12N$ and $C_{eval}^s = 12N + N^2$ are the computational complexity of evaluating the cost functions in (2.21) and (4.6), respectively, in FLOPS. As equation (4.7) shows, Δ_C increases linearly with N , exceeding 300% when $N > 36$. Since the total complexity of a minimization is dominated by the complexity of the cost function evaluations [52], equation (4.7) can be interpreted as increases in total minimization complexity.

4.3 Effects of Correlated Shadowing to Localization Performance

The results of the previous section motivate several questions about the effects of correlated shadowing. First, what is the effect of maintaining the independent shadowing assumption and using equation (2.21) to perform localization? Performing localization assuming independent shadowing when in fact the shadowing is correlated represents a model mismatch where the assumed cost function to be minimized does not match the actual phenomena that is captured in the measurements. Since undoing the independent shadowing assumption (i.e., using (4.6)) could be costly, it is important to understand how sensitive localization accuracy is to this model mismatch. Second, if the model mismatch has a significant negative effect on localization accuracy, how much of that loss can be recovered by the use of (4.6)? In this section, we address these questions by discussing prior work in this area and directly exploring the effects on the LCIS algorithm.

4.3.1 Insights from Prior Work on the Effects of Correlated Shadowing on RSS Localization

The effects of correlated shadowing on various aspects of WSN and CRN systems have been a source of attention in recent research [72–74]. In particular, localization is a major topic of interest in WSN research [72]. In [72], the authors discuss the effects of correlated shadowing on the Cramer-Rao lower bound (CRLB) on localization variance for the single transmitter case and discuss specific results for a 16-node network. In computing the

CRLB, the authors define the Fisher information matrix (FIM) for the problem as having two components: a mean component and a covariance component, both of which are inversely proportional to the CRLB. Their results indicate that on average, correlated shadowing may actually decrease the CRLB as it causes an increase in the covariance term of the FIM that is greater than the decrease it causes to the mean term. These results suggest that localization could be performed more accurately under correlated shadowing than under independent shadowing if the information in the covariance component of the FIM is utilized for estimation. A modest amount of research has been performed on localization algorithms that account for correlated shadowing in an attempt to realize the potential gains achievable if this additional covariance term information is used for localization [75, 76]. In [75], the authors propose a ML-based approach to single transmitter localization that incorporates correlated shadowing by assuming a cost function based on (4.6). The authors stress that in order to obtain the best localization results, the shadowing noise covariance matrix Σ has to be known, an assumption that may be unrealistic as Σ varies with the propagation environment and the sensor node arrangement. The authors, however, do not justify the increases in computational complexity nor put in context any performance gains achieved by the proposed algorithm by comparing its performance to that of the mismatched ML approach, in which the independent shadowing assumption is maintained. In [76], the authors propose a single transmitter localization algorithm based on differential RSS (DRSS) measurements, where the difference between RSS measurement pairs is used as the observation on which localization is based. This approach exploits the correlation among the measurements by taking advantage of the fact that the variance of a DRSS measurement decreases with increases in the correlation coefficient of the RSS measurement pair. The authors compare the localization performance of the DRSS techniques with an RSS equivalent that does not account for correlated shadowing. The results obtained in [76] show that the accuracy of the DRSS approach is higher than that of the mismatched RSS when all the measurements are highly correlated (i.e., $\rho \approx 0.8$). However, the performance of both approaches is equivalent when measurements are marginally correlated (i.e., $\rho \approx 0.4$), and

the RSS approach surpasses the DRSS as the measurements become uncorrelated. These results suggest that the DRSS (and possibly accounting for correlated shadowing itself) may not be advantageous in real life scenarios where the correlation between measurements varies in magnitude depending on the relative location of receivers.

The results discussed above, although for a single transmitter case, provide valuable insight into the general effects of correlated shadowing on the RSS localization problem. The results suggest that 1) strong correlation among all measurements is needed in order to extract enough information to significantly improve localization performance, and 2) if using an ML-based approach, Σ should be estimated accurately in order to effectively account for the correlation among measurements. Thus far, the research performed in this area does not provide conclusive results that would justify accounting for correlated shadowing in MTL. Since performing the tasks associated with accounting for this correlated shadowing may add significant computational complexity to an approach that is already relatively complex, it is important to characterize the effects of shadowing model mismatch to determine whether its severity warrants such an approach. In the next section, we explore the effects of shadowing model mismatch on the accuracy of the LCIS algorithm to better understand the potential trade offs between computational complexity and immunity to shadowing model mismatch.

4.3.2 Effects of Shadowing Model Mismatch on the Localization Accuracy of the LCIS

In this section, we explore the effects of shadowing model mismatch on the accuracy of the LCIS algorithm. We use numerical simulation to assess these effects under various propagation conditions. Simulations follow the same general setup described in Section 2.6, except that the size area of interest is varied to reflect typical propagation distances in each type of environments. The propagation conditions are parameterized by γ , σ , D_c , and the length of the area of interest L_a . Two propagations environments are simulated: 1) dense urban, where $\gamma = 4$, $\sigma = 8$, $D_c = 20$ and $L_a = 1000$; and 2) suburban, where $\gamma = 3$, $\sigma = 6$, $D_c = 500$ and $L_a = 10000$.

Figures 4.1 and 4.2 show the simulated mean performance of the LCIS algorithm under both propagation environments when shadowing is correlated and when shadowing is independent. As the figures show, the effects of shadowing model mismatch on localization accuracy appear to be minimal in both propagation environments. In the urban environment, whose performance is shown in Figure 4.1, the localization performance of the LCIS under correlated shadowing conditions (mismatched model case) does not appear to differ from its performance under independent shadowing (matched model case). This is believed to be due to the relatively low correlation among measurements (low value of D_c), which is a consequence of the diversity of clutter and obstructions that radio waves encounter in such an environment. However, in the suburban environment, whose performance is shown in Figure 4.2, the localization accuracy of the mismatched case appears to be higher than that of the matched case for $N \leq 22$. This observable is in agreement with the results of [72] and is believed to be due to a reduction of uncertainty caused by the additional information provided by the correlation among sensor measurements. This information is significant at lower values of N where the uncertainty is high, but becomes insignificant as N increases and the uncertainty is lower. The performance gap between the matched and unmatched cases reverses at values of $N \geq 22$, where the performance of the matched case becomes greater. This reverse is believed to be due to a rise in uncertainty caused by a reduction in the total information contained in measurements (i.e., for large N , the decrease of the mean term of the FIM is greater than the increase of the covariance term). This observed behavior suggests that the results of [72] either do not apply to all values of N , or that they do not manifest themselves in the same manner when interference is present.

The results from our simulations suggest that the effects of shadowing model mismatch on MTL accuracy do not warrant the costs and complexities associated with pursuing a matched cost function approach. As discussed above, possible match model approaches would employ cost function (4.6) and an algorithm to estimate Σ in order to avoid further model mismatches caused by using an inaccurate Σ . These approaches represent an increase in complexity of at least 300% for $N \geq 36$. Finally, the results confirm the validity of the

independent shadowing assumption for ML-based MTL.

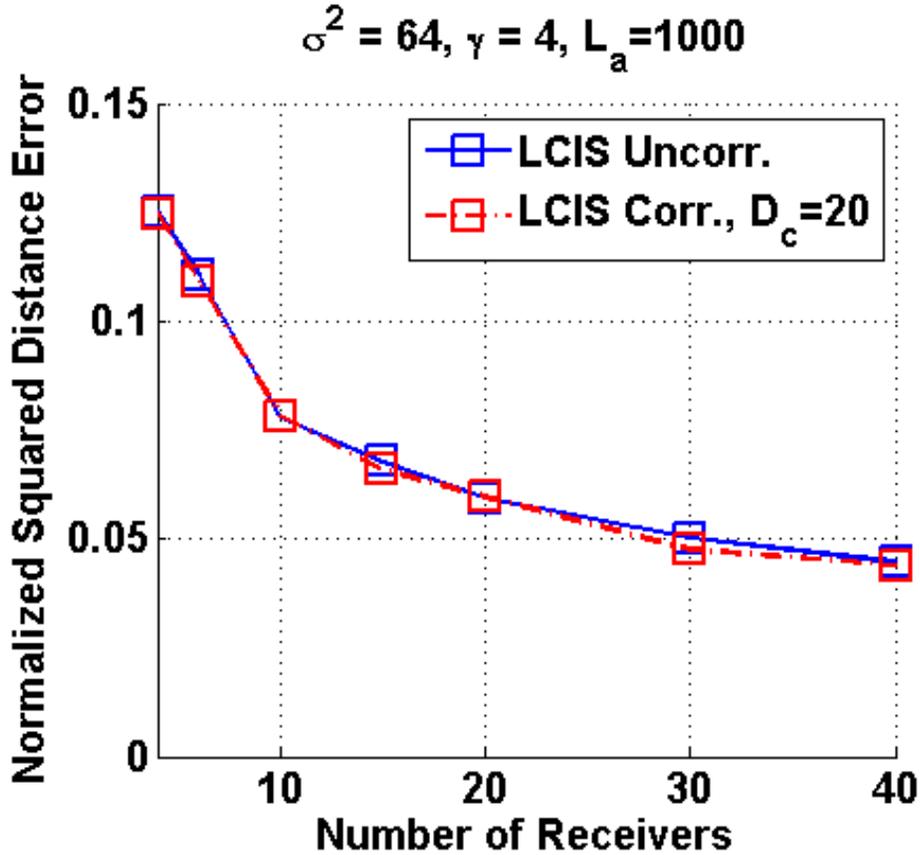


Figure 4.1: Two-transmitter mean performance comparison of LCIS algorithm (assuming uncorrelated shadowing) under uncorrelated and correlated shadowing operating in a dense urban propagation environment with parameters $\sigma^2 = 64, \gamma = 4, D_c = 20$ and $L_a = 1000$.

4.4 Exploitation of Mobility in MTL

As sensor nodes move to different locations, they have the opportunity to take additional power measurements at these new locations. These new measurements contain additional information about the transmitter locations. In this section, we explore ways of extracting this extra information in an efficient manner, such that localization accuracy can be improved without wasting valuable computational resources. We employ the insight gained

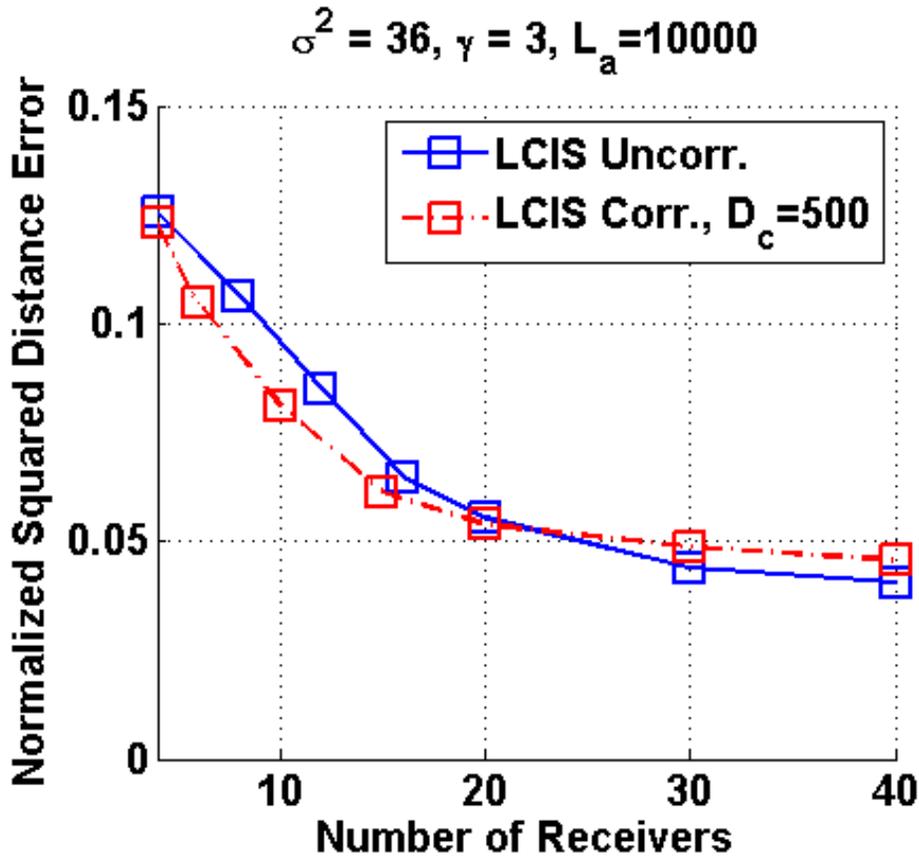


Figure 4.2: Two-transmitter mean performance comparison of LCIS algorithm (assuming uncorrelated shadowing) under uncorrelated and correlated shadowing operating in a sub-urban propagation environment with parameters $\sigma^2 = 36$, $\gamma = 3$, $D_c = 500$ and $L_a = 10000$.

in the interference analysis of Chapter 2 to devise a strategy to reduce overall measurement uncertainty and to develop an extension to the LCIS algorithm that enables it to take advantage of the newly available information.

4.4.1 Extended LCIS for Mobility Exploitation

One major finding in exploring the MTL problem has been the establishment of a relationship between sensor-transmitter configuration and localization accuracy. As discussed in Chapter 2, the expected estimation performance is a function of the interference noise statistics, which are different for each receiver and are in turn a function of each receiver's

distance ratio (or distance-power ratio when transmitted powers are unequal). The localization algorithms developed thus far operate on the principle of adjusting the observations to remove the expected interference. This technique has proven to be effective; however, one interesting observation that could lead to possible improvements is that depending on sensor-transmitter configuration, the algorithms may use measurements with high interference noise variance. These high uncertainty measurements can considerably degrade estimation accuracy. In view of this, we consider ways of incorporating network configuration information into the estimation process with the objective of reducing the overall measurement uncertainty (interference variance), which results in improved estimation performance. We explore two techniques that aim to achieve this goal by using the distance ratio as a metric of measurement uncertainty. The first technique is distance-ratio-based clustering, in which distance ratio is used to determine to which cluster a measurement is assigned. As it will be shown, this technique bounds distance ratio to a maximum value equal to 1, thus reducing overall measurement uncertainty. The second technique we explore is distance-ratio-based measurement discrimination. This technique calls for requiring that the measurements used for the estimation of a particular transmitter are taken by sensors with distance ratios smaller than a predefined threshold. This action further bounds the expected interference variance that any measurement could exhibit, resulting in a configuration with less measurement uncertainty. This technique focuses on achieving higher algorithmic efficiency rather than on increasing localization accuracy given a fixed value of N (i.e., achieving the biggest accuracy gains with the smallest increase in computational complexity). This idea is centered on the fact that measurements with higher uncertainty carry less information than those with lower uncertainty, and thus contribute less to the estimation solution.

We propose an extended low-complexity IS algorithm where we use estimates of distance ratio to measure the uncertainty of individual measurements and to implement the techniques described above. To maintain low computational complexity, the LCIS is extended to refine previously computed estimates without executing an additional full run of

the algorithm. The extensions include distance ratio estimation, re-clustering, and refining minimizations. As will be discussed later in the section, these extensions represent a complexity increase of approximately 50% to 100% over the complexity of the LCIS algorithm. The proposed extended low-complexity IS algorithm is described below.

Extended low-complexity IS algorithm

1. Run an instance of the LCIS algorithm as described in Section 2.3.2 using N measurements taken at initial receiver locations $\Phi^{(0)}$.
2. Let receivers move to new locations denoted by $\Phi^{(1)}$ in accordance with the RWM described in Section 4.1.1. Concatenate measurements from both sets of locations as

$$\mathbf{r} = \begin{bmatrix} \mathbf{r}^{(0)} \\ \mathbf{r}_{(dB)}^{(0)} \end{bmatrix} \mathbf{r}^{(1)} \mathbf{r}_{(dB)}^{(1)}. \quad (4.8)$$

3. Cluster measurements in \mathbf{r} based on their proximity to estimates $\hat{\theta}^{(0)}$ obtained on first run of LCIS. Clustering is performed as follows. Let the set $\mathcal{S} = \{l\}_{l=1}^{2N}$ be the set containing the indices of all measurements and \mathcal{C}_i be the set containing the indices of the measurements in the i th cluster, $i \in [1, 2, \dots, M]$. Cluster \mathcal{C}_i is defined as

$$\mathcal{C}_i = \left\{ l : \min_k \{d_{kl}\} = d_{il}, k \in [1, 2, \dots, M], \forall l \in [1, 2, \dots, 2N] \right\}. \quad (4.9)$$

4. Compute estimates of the distance ratios as

$$\frac{\hat{d}_{ml}}{\hat{d}_{il}} = \frac{\sqrt{(\hat{x}_m^{(0)} - x_l)^2 + (\hat{y}_m^{(0)} - y_l)^2}}{\sqrt{(\hat{x}_i^{(0)} - x_l)^2 + (\hat{y}_i^{(0)} - y_l)^2}}, \quad (4.10)$$

where $m \in [1, 2, \dots, M]$, $i \in [1, 2, \dots, M]$, $i \neq m$, $l \in [1, 2, \dots, 2N]$, and $(\hat{x}_i^{(0)},$

$\hat{y}_i^{(0)}$ is an estimate of the location of the i th transmitter generated using measurements taken at the initial set of receiver locations, and (x_l, y_l) is the location where the l th measurement was taken.

5. Compare $\frac{\hat{d}_{ml}}{\hat{d}_{il}}$ to threshold T_g and classify the l th measurement as having high uncertainty if

$$T_g < \frac{\hat{d}_{ml}}{\hat{d}_{il}}, \quad (4.11)$$

where $0 < T_g \leq 1$; $m \in [1, 2, \dots, M]$, $i \in [1, 2, \dots, M]$, $i \neq m$; and $l \in [1, 2, \dots, 2N]$. The measurements classified as having high uncertainty are not used for estimating a new set of location estimates. The total number of receivers in the resulting set of clusters is defined as N_r .

6. Perform Steps 3, 4 and 5 of the LCIS algorithm as described in Section 2.3.2 to produce a refined and final estimate

As the enumeration shows, the algorithm starts by computing an initial set of estimates $\hat{\theta}^{(0)}$ using measurements $\mathbf{r}_{(dB)}^{(0)}$ taken at initial locations $\Phi^{(0)}$. In Step 2, new measurements are collected at new receiver locations $\Phi^{(1)}$ and concatenated with the previous set of measurements to form (4.8). In Step 3, the vector of measurements \mathbf{r} is divided into clusters based on each receiver's proximity to $\hat{\theta}^{(0)}$. In this step, the information contained in the initial transmitter location estimates is used to achieve a clustering that reduces the overall measurement uncertainty by assigning each receiver to the cluster that results in lowest estimated distance ratio. In addition, since each receiver is assigned to the cluster estimating the location of the transmitter that is believed to be closest to it, the numerator in its estimated distance ratio will always be less than 1, thus bounding the measurement uncertainty in each cluster. This is the main vehicle for achieving localization performance; by lowering the per-measurement uncertainty, the number of receivers per cluster that have acceptable levels of uncertainty is maximized. In Steps 4 and 5, the clusters are refined

by removing measurements that are expected to contain high levels of uncertainty. This is achieved by using the estimated distance ratio as a predictor of measurement uncertainty and removing measurements that exceed the distance ratio threshold T_g . This threshold is set according to how much uncertainty can be tolerated. As Figure 2.5 shows, the variance of the interference noise varies with the propagation conditions. Thus, a lower threshold may be desired for a suburban/rural environment, whereas a higher threshold may be used for an urban environment. The effects of uncertainty reduction in this step are marginal, as the majority of the uncertainty is removed by the proximity-based clustering of Step 3, which limits distance ratios to a maximum value of 1. The main advantage of removing measurements is that it saves computational resources, as the measurements with distance ratio close to 1 do not contribute to estimation accuracy as much as others with values significantly lower than 1. The last step corresponds to the interference subtraction and estimation steps of the LCIS algorithm. In this step, the initial estimates $\hat{\theta}^{(0)}$ are used to subtract the expected interference from the received power measurements. Since, on average, the location estimates $\hat{\theta}^{(0)}$ are more accurate than those generated internally within the LCIS, the second interference subtraction step of the extended LCIS algorithm is more effective at removing the expected interference and thus provides a set of measurements that better match the assumed likelihood function. Once the expected interference is reduced, a new set of estimates $\hat{\theta}^{(1)}$ are computed by solving (2.24) and a final set of estimates is selected according to (2.25).

4.4.2 Simulation Results

In this section, the localization performance of the extended LCIS algorithm is compared to that of the LCIS. The objective of the comparison is to determine the performance gains obtained by using additional measurements taken at different locations. The comparison is performed via numerical simulations.

Simulations follow the same general setup described in Section 2.6, where the transmitters are assumed to be separated by at least 20% of the length of the area of interest.

Figure 4.3 shows the simulated performance gain of the extended LCIS relative to that of the LCIS algorithm for different values of the distance ratio threshold T_g . As with previous localization performance simulations, the chosen performance metric is the average squared distance error between estimated and true transmitter locations, where the average is taken over the M transmitter location estimates. The figure shows the relative mean error over 1000 different random draws of $M = 2$, and $N = 4$ to $N = 40$ receivers with shadowing variance of $\sigma^2 = 36$ and $\gamma = 3$. Localization performance gain curves are plotted for distance ratio thresholds $T_g = 0.65, 0.75, 0.85$ and 1 , where each curve is comprised of a set of data points and a dashed line. The dashed lines represent a least squares (LS) first order

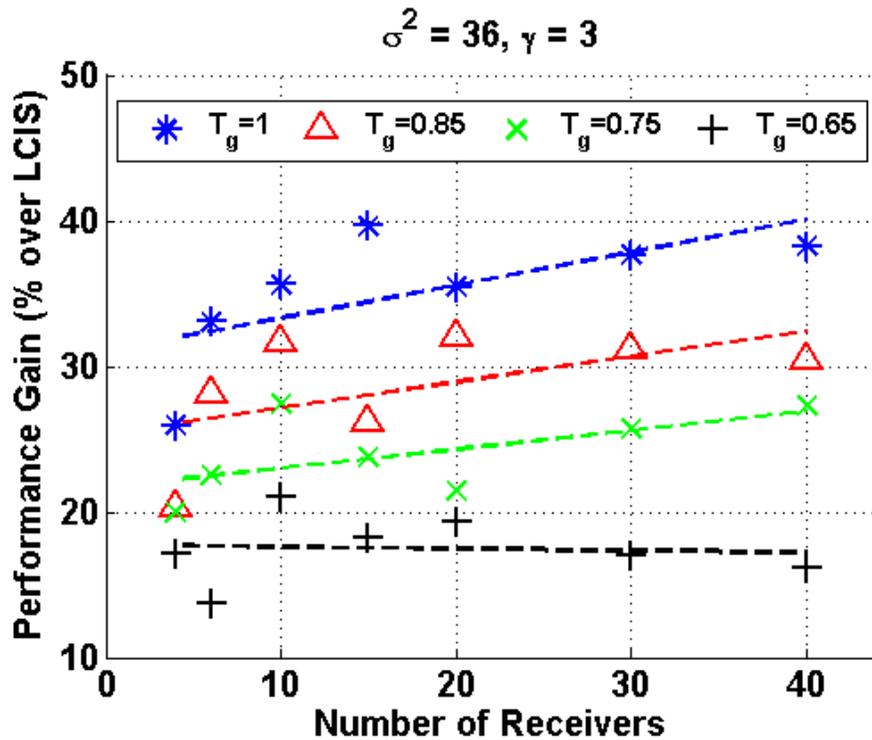


Figure 4.3: Two-transmitter performance gain as a percentage of LCIS accuracy versus the number of receivers for various distance ratio thresholds in a propagation environment with $\sigma^2 = 36$ and $\gamma = 3$.

polynomial fit over the data points, which is done to reduce variability in data in order to better visualize the trends each curve has with N . As the plot shows, performance gains increase relatively uniformly with T_g with a maximum increase of approximately 40% for

$T_g = 1$ and a minimum of approximately 18% for $T_g = 0.65$. This is believed to be due to the fact that as T_g increases, more measurements are used for estimation, providing the estimator with additional information. As the first order fits to the data show, localization performance gains also increase with N except for when $T_g = 0.65$, where the gains appear relatively constant. In this particular case, the number of receivers used for estimation is approximately equal to N , showing that the developed refining technique achieves gains even when $N_r = N$. To better illustrate the relationship between the number of receivers used for refinement and performance gains, Figures 4.4 and 4.5 show contour plots of the performance gains versus the number of receivers and the percentage of the total number of available measurements used for refinement in two different propagation environments. The percentage of the total number of available measurements used for refinement is computed as $\mathcal{E}[N_r]/2N$. As seen in the contours, performance gains increase with $\mathcal{E}[N_r]$ for all values of N , suggesting that even measurements with the highest distance ratios (near 1) contribute positively to mean localization accuracy. The contours also show that higher $\mathcal{E}[N_r]$ values result in larger performance increases at the largest values of N , with the largest increases occurring at $N = 40$. At the lower N values, however, the performance gains are still considerable, with a maximum performance gain of approximately 32% for $N = 4$. The plots also show that performance gains of approximately 17% are achieved even when the percentage of total measurements used for refinement is 50% ($N_r = N$), suggesting that both measurement re-clustering and elimination based on distance ratio do reduce the overall measurement uncertainty.

4.5 Incorporating Refining Technique into the LCIS Framework

The work of the previous sections revealed that the accuracy of the set of initial estimates is key in reducing measurement uncertainty and refining transmitter location estimates. Results suggests that this lessons may be incorporated into the original IS framework to

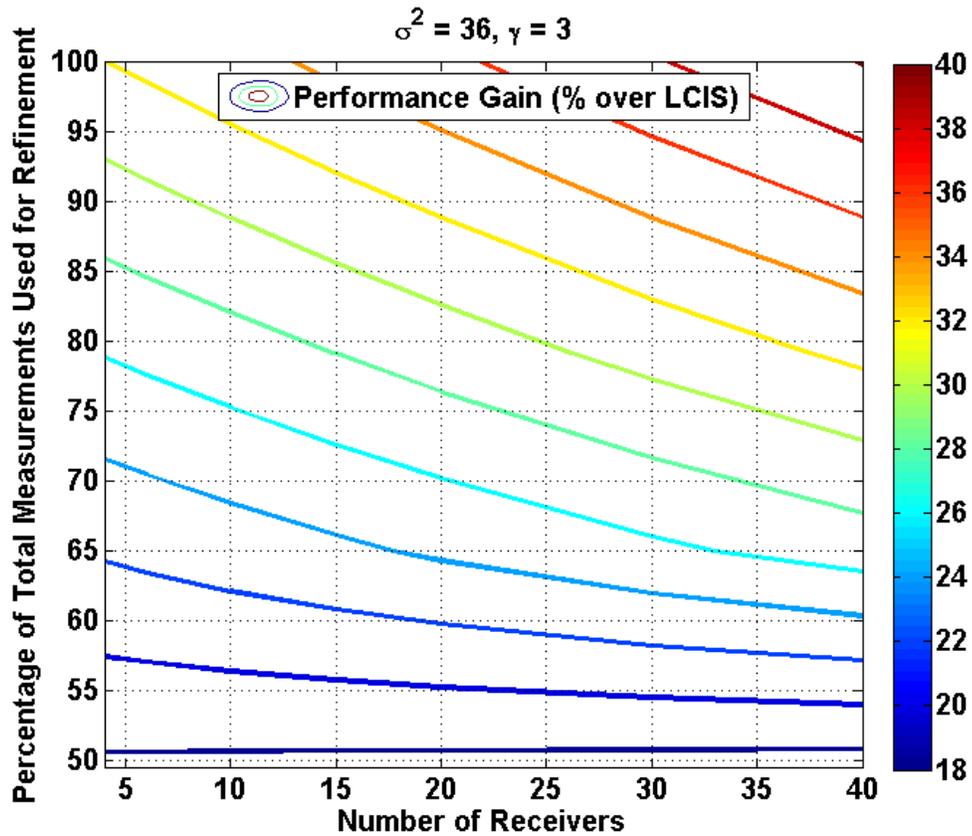


Figure 4.4: Two-transmitter performance gain as a percentage of LCIS accuracy versus the number of receivers and percentage of total measurements used for estimate refinement for various distance ratio thresholds in a propagation environment with $\sigma^2 = 36$ and $\gamma = 3$.

obtain performance gains while maintaining or reducing its computational complexity. In this section, we study the computational complexity of the extended LCIS to motivate the incorporation of its refining technique into the LCIS framework. Additionally, to show how the performance of the LCIS is improved, we compare the simulated localization accuracy of the LCIS to that of the extended LCIS with equivalent computational complexity.

As detailed in Section 3.6.1, the computational complexity of the LCIS algorithm is measured based on the number of nonlinear minimizations. The computational complexity of a single minimization has two main components: 1) initialization and termination tests, and 2) evaluations of the cost function [52]. The initialization and termination tests component is constant with the number of receivers N , while the cost function evaluations

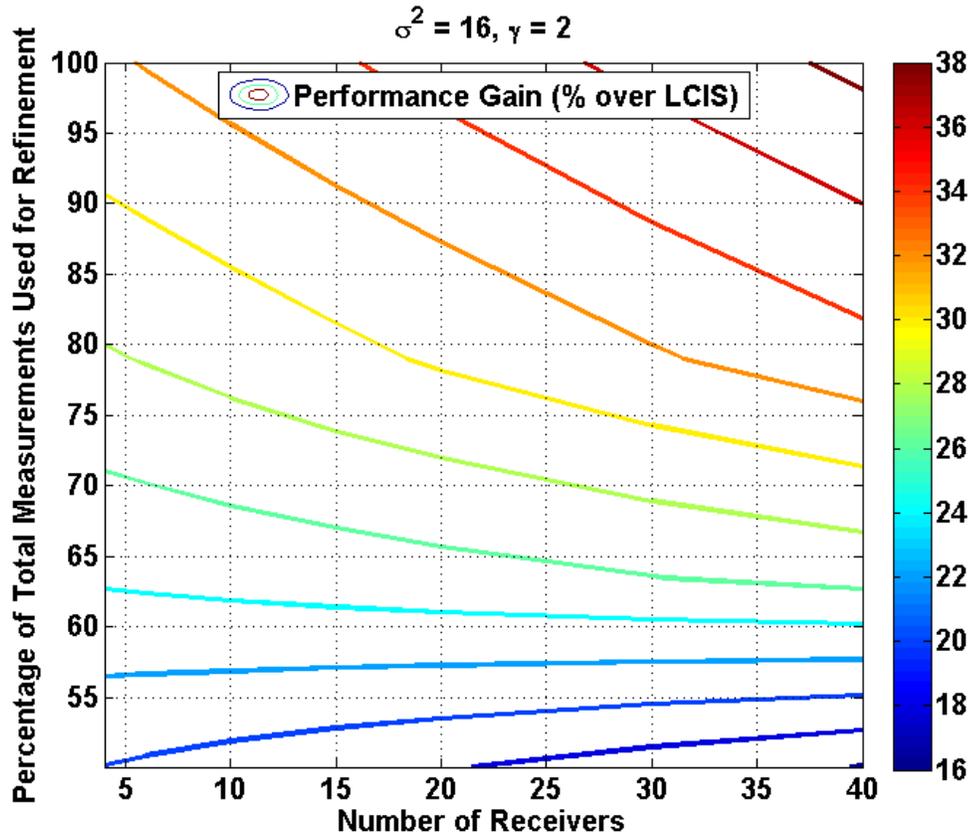


Figure 4.5: Two-transmitter performance gain as a percentage of LCIS accuracy versus the number of receivers and percentage of total measurements used for estimate refinement for various distance ratio thresholds in a propagation environment with $\sigma^2 = 16$ and $\gamma = 2$.

component is a monotonically increasing function of N . Moreover, as shown in Figure 3.8, as N increases, the cost function evaluations dominate the total computational complexity. Thus, the total computational complexity of the minimizations can be approximated by its cost function evaluation component. Recall (3.45), which states this relation as

$$C_{nm}(N) \approx 2I_{nm}C_{eval}(N), \quad (4.12)$$

where I_{nm} is the number of NM iterations required for convergence, and $C_{eval}(N)$ is the complexity of a single cost function evaluation. Since for the cost function (3.43),

$C_{eval}(N) = 12N$, a linear function with y -intercept equal to 0, we can define the computational complexity of a minimization using aN receivers as

$$\begin{aligned} C_{nm}(aN) &\approx 2I_{nm}C_{eval}(aN) \\ &= 2aI_{nm}C_{eval}(N) \\ &= aC_{nm}(N), \end{aligned} \tag{4.13}$$

where a is a real positive number. This result facilitates the expression of the complexity of a minimization that uses aN observations as a fraction of a minimization that uses N observations. Recall that the LCIS complexity is given by

$$C_L(N) = 2MC_{mn}(N). \tag{4.14}$$

The extended LCIS has a complexity given by

$$\begin{aligned} C_{Lext}(N) &= C_L(N) + MC_{mn}(N_r) \\ &= 2MC_{mn}(N) + MC_{mn}(bN) \\ &= (2 + b)MC_{mn}(N), \end{aligned} \tag{4.15}$$

where $b \in (0, 2]$ describes the number of measurements used for refinement as a multiple of N . As (4.15) shows, the complexity of the extended LCIS has two components: 1) the computations spent by the LCIS algorithm to generate the first set of estimates, and 2) the computations spent by M minimization using N_r measurements, which are performed to refine the estimates produced by the LCIS. The equation also shows that the extended LCIS algorithm can have up to twice the computational complexity of the LCIS, quantifying how much additional complexity is required to achieve the performance gains demonstrated in Section 4.4.2. It is worth mentioning that in the case where $b = 2$ where $C_{Lext}(N) = C_L(2N)$, the extended LCIS algorithm performs better than a single LCIS run

with $2N$ measurements. This is believed to be due to the fact that the final estimates out of the LCIS using N measurements provide a more accurate set of initial estimates for interference subtraction than the LCIS's internal initial estimates using $2N$ measurements. This rationale is what motivates incorporation of the refinement methodology used in the extended LCIS algorithm into the LCIS algorithm. Such change can be implemented by running an instance of LCIS using half of the N measurements available for a single run of the algorithm, and then refining the estimates generated by the LCIS run according to the refining procedure of the extended LCIS using all N measurements. This algorithm would have the following computational complexity:

$$\begin{aligned}
 C_{Lmod}(N) &= C_L(N/2) + MC_{mn}(N) & (4.16) \\
 &= MC_{mn}(N) + MC_{mn}(N) \\
 &= 2MC_{mn}(N) \\
 &= C_L(N).
 \end{aligned}$$

As (4.16) shows, the complexity of such an algorithm would be exactly the same as that of the LCIS algorithm. Figure 4.6 shows a plot of the expected localization performance of the modified LCIS compared to that of the LCIS. Since the extended LCIS algorithm using $N_r = 2N$ measurements for refinement is equivalent to the modified LCIS using N measurements, the performance curve for the modified LCIS was obtained by expanding the results of the extended LCIS by a factor of 2 as denoted by

$$E_{\theta Lmod}(N) = E_{\theta Lext}(2N). \quad (4.17)$$

As seen in the figure the expected localization accuracy of the modified LCIS is essentially equal to that of the LCIS for the lower N values, which is believed to be due to the fact that at low values of N , $N/2$ measurements are not enough to provide a sufficiently accurate initial estimate for refinement to provide gains. However, as N increases, the performance

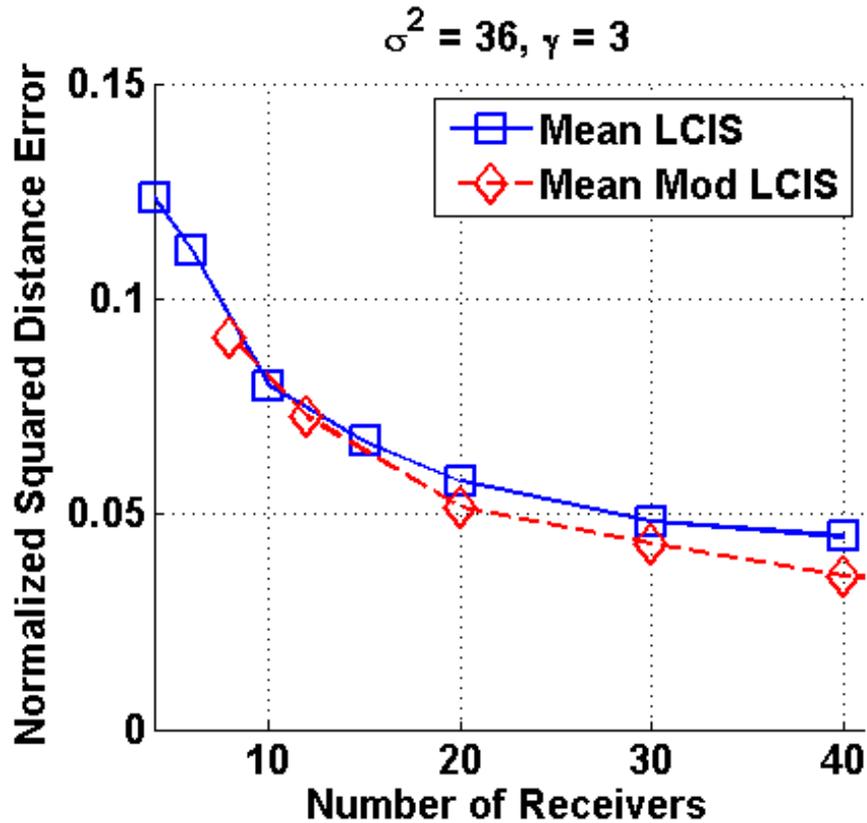


Figure 4.6: Two-transmitter mean localization accuracy of the LCIS and modified LCIS versus the number of receivers under $\sigma^2 = 36$ and $\gamma = 3$.

gap between the two algorithms widens, with the modified LCIS providing a 22% increase in accuracy at $N = 40$. Another way to interpret the modified LCIS curve in Figure 4.6 is to view it as the performance of the extended LCIS with $T_g = 1$ normalized by the maximum number of measurements used. Analyzing the results from this viewpoint suggests that the proposed clustering and refining technique increases the accuracy per measurement used for estimation, with the gains in accuracy increasing with N .

Chapter 5: Conclusion and Future Work

5.1 Conclusion

Due to its simplicity and low cost of implementation, RSS-based localization is an attractive localization technique for resource-constrained systems such as CRN's and WSN's. In these systems, measurements at sensor nodes often contain co-channel interference. This interference significantly reduces localization accuracy and complicates the localization task. One of the biggest challenges in solving this problem is that there exists no closed-form probability distribution function for the observations made at the sensor nodes. To address this challenge, a novel interference analysis was developed in which a closed-form approximation to the density function of the interference affecting RSS measurements was derived. Closed-form expressions for the mean and variance of the interference were also derived. A key finding of the analysis is that for a given propagation environment, the statistics of the interference at a particular node are strictly a function of the distance ratio parameter. Analysis results lead to the development of an estimation approach called interference subtraction (IS). The IS approach consists of subtracting the expected interference from each receiver's observation, a process that allows the MTL problem to be broken into multiple single transmitter localizations. In addition, IS enables the estimation of multiple transmitter locations using estimators for the single transmitter case, which are typically simpler to derive. We have shown how the IS approach can be applied to different estimation techniques by developing two MTL algorithms, the LCIS algorithm and the IS-based EM algorithm, both of which employ the single transmitter maximum likelihood estimator. The LCIS algorithm was shown to outperform competing algorithms by as much as 3 dB when computational complexity is constrained. On the other hand, the IS-based EM was shown to equal the performance of the competing algorithms, while offering an alternative

derivation to the EM approximation for solving the MTL problem. The developed algorithms span the computational complexity versus localization accuracy trade off and offer system designers high fidelity within the trade.

To avoid the computational challenges that encumber ML-based estimation techniques, we have studied the application of a linear estimation approach to MTL. We have also studied the problem of distributing localization computations over the sensor nodes. We have extended and applied prior work on information dissemination and linear single transmitter localization to develop a linear distributed MTL algorithm, namely the MTWARL algorithm. The algorithm uses a linear, closed-form, proximity-based location estimator, the weighted average receiver location, to iteratively cluster and estimate transmitter locations. The algorithm employs a randomized pair-wise gossip algorithm to perform computations in a distributed fashion. It has been shown that for large number of sensor nodes, the MTWARL algorithm can achieve greater localization accuracy than that of the IS-based centralized algorithms, while using as little as half of the energy in heavily obstructed communication environments. The convergence properties of the randomized pair-wise gossip algorithm have been studied, and the configuration that minimizes the lower bound on convergence time has been derived. This result is critical to reducing overall energy utilization of the distributed MTL algorithm. Concluding the work on this topic, comparative analysis has been performed in which the localization accuracy, energy utilization, and computational complexity of the developed distributed algorithm have been compared to that of the competing centralized algorithms. This analysis enables the study of system-level trades related to the implementation of MTL algorithms.

As a final extension of our work on MTL, the effects of sensor mobility on the statistical properties of RSS observations and its implications for ML-based MTL have been studied. The major implication was found to be a change of the cost function induced by the introduction of correlated shadowing into the system model. Most of the literature on this topic assumes that these changes must be dealt with and that alternative approaches must be pursued in order to avoid and/or take advantage of these effects. Our studies have revealed

that the effects of correlated shadowing to MTL are minimal and that the gains obtained by additional processing to account for correlated shadowing are too small to justify large increases in computational complexity. As part of this work, the LCIS algorithm has been extended to take advantage of additional measurements enabled by sensor mobility. It was demonstrated that additional measurements can be used to iteratively reduce overall measurement uncertainty and thus improve localization performance.

5.2 Future Work

5.2.1 Algorithms for Estimating the Number of Transmitters

The MTL framework developed in this dissertation assumes that the number of simultaneous emitters M is known. In practice, depending on the application, this assumption may not be valid. In these cases, the assumption of known M relies on the use of external algorithms for the estimation of M prior to estimating transmitter locations. These algorithms are often highly complex, and thus require significant computational resources for their implementation. To that end, an important area of future study is the development of simple estimation techniques for the number of transmitters M . Recent work on topological signal processing suggests that employing homology, simple algorithms can be derived for this purpose in a coordinated measurement system. However, it is not clear whether such techniques would be applicable to MTL, where the measurements are taken in an uncoordinated fashion.

5.2.2 Alternative Gossip Algorithms for MTL

The MTWARL was shown to offer relatively good performance while offering energy savings over the centralized approaches in high path loss exponent environments. However, the energy utilization in low path loss exponent environments is still quite high compared to that of the centralized approaches. To expand the appeal of the MTWARL, it would be desirable to lower the energy utilization when operating in these environments. The type of

gossip algorithm highly influences the energy utilization. Pair-wise gossip is simple and is guaranteed to converge. However, its convergence time is not as fast as other types of gossip algorithms that may not offer the same convergence properties and simplicity. An area of future study is the use of faster gossip algorithms and explore the trade offs among energy utilization, complexity, and stability. In particular, broadcast gossip algorithms would be of interest, as they employ a one-to-many updating mechanism that may reduce convergence times.

Bibliography

Bibliography

- [1] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379–423 and 623–656, July and October 1948.
- [2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near shannon limit error-correcting coding and encoding: Turbo-codes," in *Proc. of the IEEE International Conference on Communications 1993*, May 1993.
- [3] FCC Spectrum Policy Task Force, "Report of the spectrum efficiency working group," Tech. Rep., Nov. 2002.
- [4] Q. Zhao and B. Sadler, "A survey of dynamic spectrum access," *IEEE Signal Processing Magazine*, vol. 24, no. 3, May 2007.
- [5] D. Hatfield and P. Weiser, "Property rights in spectrum: Taking the next step," in *IEEE DYSpan Symp.*, Nov. 2005, pp. 43–55.
- [6] FCC, "Title 47 of the code of federal regulations (CFR), Part 15," [Online], <http://www.fcc.gov/oet/info/rules/>.
- [7] R. Etkin, A. Parekh, and D. Tse, "Spectrum sharing for unlicensed bands," *IEEE Journal on Selected Areas in Comm.*, vol. 25, no. 3, pp. 517–528, Apr. 2007.
- [8] K. Gilhousen, I. Jacobs, R. Padovani, A. Viterbi, J. Weaver, L.A., and I. Wheatley, C.E., "On the capacity of a cellular CDMA system," *Vehicular Technology, IEEE Transactions on*, vol. 40, no. 2, pp. 303–312, May 1991.
- [9] M. Win and R. Scholtz, "Impulse radio: how it works," *Communications Letters, IEEE*, vol. 2, no. 2, pp. 36–38, Feb. 1998.
- [10] J. Mitola, "Cognitive radios for flexible mobile multimedia communications," in *Proc. IEEE International Workshop on Mobile Multimedia Communications*, Oct. 1999.
- [11] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE Journal on Selected Areas in Comm.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [12] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys Tutorials, IEEE*, vol. 11, no. 1, pp. 116–130, 2009.
- [13] R. Menon, R. M. Buehrer, and J. H. Reed, "Outage probability based comparison of underlay and overlay spectrum sharing techniques," in *Proc. of IEEE DySPAN*, Nov. 2005.

- [14] J. K. Nelson, M. Hazen, and M. R. Gupta, "Global optimization for multiple transmitter localization," in *Proc. of Military Communications Conf. (MILCOM)*, Oct. 2006.
- [15] B. Mark and A. Nasif, "Estimation of interference-free transmit power for opportunistic spectrum access," in *Proceedings of the IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2008.
- [16] "Evaluation of the performance of prototype TV-band white space devices phase ii: Report fcc/oet 08-tr-1005," FCC OET, Tech. Rep., 2008.
- [17] D. Gurney, G. Buchwald, L. Ecklund, S. Kuffner, and J. Grosspietsch, "Geo-location database techniques for incumbent protection in the TV white space," in *Proc. of IEEE DySPAN*, Oct. 2008.
- [18] N. Patwari, A. O. H. III, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, no. 8, pp. 2137–2148, Aug. 2003.
- [19] X. Li, "Collaborative localization with received-signal strength in wireless sensor networks," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, Nov. 2007.
- [20] H. Wymeersch, J. Lien, and M. Win, "Cooperative localization in wireless networks," *Proceedings of the IEEE*, vol. 97, no. 2, Feb. 2005.
- [21] N. Patwari, J. N. Ash, S. Kyperountas, A. O. H. III, R. L. Moses, and N. S. Correal, "Locating the nodes," *IEEE Signal Processing Magazine*, vol. 22, no. 4, Jul. 2005.
- [22] A. Nasif and B. Mark, "Collaborative opportunistic spectrum sharing in the presence of multiple transmitters," in *Proc. IEEE Globecom'08*, Dec. 2008.
- [23] M. Hata, "Empirical formula for propagation loss in land mobile radio services," *Vehicular Technology, IEEE Transactions on*, vol. 29, no. 3, pp. 317 – 325, Aug. 1980.
- [24] Y. Okumura, E. Ohmori, T. Kawano, and K. Fukuda, "Field strength and its variability in VHF and UHF land-mobile radio service," *The Electrical Communication Laboratory, Review of*, vol. 16, pp. 835–873, Sept. 1968.
- [25] A. Goldsmith, *Wireless Communications*. New York, NY: Cambridge University Press, 2005.
- [26] A. J. Coulson, A. G. Williamson, and R. G. Vaughan, "A statistical basis for lognormal shadowing effects in multipath fading channels," *IEEE Transactions on Vehicular Technology*, vol. 46, no. 4, pp. 494–502, Apr. 1998.
- [27] V. Erceg, L. Greenstein, S. Tjandra, S. Parkoff, A. Gupta, B. Kulic, A. Julius, and R. Bianchi, "An empirically based path loss model for wireless channels in suburban environments," *Selected Areas in Communications, IEEE Journal on*, vol. 17, no. 7, pp. 1205 –1211, Jul. 1999.
- [28] S. Ghassemzadeh, L. Greenstein, A. Kavcic, T. Sveinsson, and V. Tarokh, "UWB indoor path loss model for residential and commercial buildings," in *Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th*, vol. 5, Oct. 2003, pp. 3115 – 3119 Vol.5.

- [29] M. Feder and E. Weinstein, "Parameter estimation of superimposed signals using the EM algorithm," *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 36, no. 4, pp. 477–489, 1988.
- [30] X. Sheng and Y. Hu, "Maximum likelihood multiple-source localization using acoustic energy measurements with wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 53, no. 1, Jan. 2005.
- [31] S. S. Reddy, "Multiple source location - a digital approach," *IEEE Trans. on Aerospace and Electronic Systems*, vol. 15, no. 1, pp. 95–105, Jan. 1979.
- [32] J. K. Nelson and M. R. Gupta, "An EM technique for multiple transmitter localization," in *Proc. of the 41st Conf. on Information Science and Systems*, 2007.
- [33] A. Dogandzic and P. Amran, "Signal-strength based localization in wireless fading channels," in *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems, and Computers*, Nov. 2004, pp. 2160–2164.
- [34] G. L. Stuber, *Principles of Mobile Communications*. New York: Springer, 1996.
- [35] S. Kim, H. Jeon, and J. Ma, "Robust localization with unknown transmission power for cognitive radio," in *Proc. of Military Communications Conf. (MILCOM)*, Oct. 2007.
- [36] J. K. Nelson, M. R. Gupta, J. E. Almodovar, and W. H. Mortensen, "A quasi EM method for estimating multiple transmitter locations," *IEEE Signal Processing Letters*, vol. 16, no. 8, pp. 354–357, 2009.
- [37] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Sixth Intl. Symp. on Micro Machine and Human Science*. IEEE, Oct. 1995, pp. 39–43.
- [38] J. K. Nelson, J. E. Almodovar, M. R. Gupta, and W. H. Mortensen, "Estimating multiple transmitter locations from power measurements at multiple receivers," in *Proceedings of the IEEE International Conference on Acoustic, Speech, and Signal Processing*, Apr. 2009.
- [39] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2nd ed. Upper Saddle River, NJ: Prentice Hall, 2001.
- [40] M. Gudmunson, "Correlation model for shadow fading in mobile radio systems," *Electronic Letters*, vol. 27, no. 23, pp. 2145–2146, 1991.
- [41] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. Boston, MA: McGraw-Hill, 1991.
- [42] N. S. Asaithambi, *Numerical Analysis: Theory and Practice*. Philadelphia, PA: Saunders, 1995.
- [43] M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. Mineola, NY: Dove, 1972.
- [44] L. F. Fenton, "The sum of lognormal probability distributions in scatter transmission systems," *IRE Trans. Commun. Syst.*, vol. CS-8, pp. 57–67, 1960.

- [45] H. L. V. Trees, *Detection, Estimation, and Modulation Theory*. New York, NY: Wiley & Sons, 2001.
- [46] G. A. F. Seber, *Multivariate Observations*. Hoboken, NJ: John Wiley and Sons, 1984.
- [47] T. Orchard and M. A. Woodbury, "A missing information principle: theory and applications," in *Proc. 6th Berkeley Symposium on Math., Statist. and Prob.*, 1972, pp. 697–715.
- [48] R. Sundberg, "Maximum likelihood theory for incomplete data from an exponential family," *Scandinavian Journal of Statistics*, no. 1, pp. 49–58, Feb. 1974.
- [49] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *Journal of the Royal Statistical Society*, vol. 39, no. 1, pp. 1–38, 1977.
- [50] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Processing Magazine*, vol. 13, no. 6, Nov. 1996.
- [51] G. McLachlan and T. Krishnan, *The EM Algorithm and Extensions*. New York: John Wiley and Sons, 1996.
- [52] S. Singer and S. Singer, "Complexity analysis of Nelder-Mead search iterations," in *Proceedings of the Conference on Applied Mathematics and Computation*, Sept. 1999.
- [53] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Oxford Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.
- [54] 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, no. 1, pp. 112–147, 1998.
- [55] J. Tsitsiklis, "Problems in decentralized decision making and computation," Ph.D. Dissertation, Lab. Information and Decision Systems, MIT, Cambridge, MA, 1984.
- [56] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah, "Randomized gossip algorithms," *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2508–2530, Jun. 2006.
- [57] A. G. Dimakis, S. Kar, J. M. Moura, M. G. Rabbat, and A. Scaglione, "Gossip algorithms for distributed signal processing," in *Proceedings of the IEEE*, Nov. 2010, pp. 1847–1864.
- [58] R. Karp, C. Schindelhauer, S. Shenker, and B. Vocking, "Randomized rumor spreading," in *Proceedings of Annual Symposium on Foundation of Computer Science*, Nov. 2000, pp. 565–574.
- [59] M. Rabbat, R. Nowak, and J. Bucklew, "Robust decentralized source localization via averaging," in *Proceedings of the IEEE International Conference on Acoustic, Speech, and Signal Processing*, Mar. 2005.
- [60] C. Laurendeau and M. Barbeau, "Centroid localization of uncooperative nodes in wireless networks using a relative span weighting method," *EURASIP Journal of Wireless Communications Networks*, Sept. 2010.

- [61] J. Blumenthal, R. Grossmann, F. Golatowski, and D. Timmermann, "Weighted centroid localization in zigbee-based sensor networks," in *Proceedings of the IEEE Int. Symp. on Intelligent Signal Processing*, Oct. 2007.
- [62] M. C. Grant, CVX Research, Austin, TX, "CVX," <http://cvxr.com/>.
- [63] J. F. Sturm, McMaster University, Ontario, Canada, "SeDuMi: Convex optimization solver," <http://sedumi.ie.lehigh.edu/>.
- [64] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.
- [65] J. Markoulidakis, G. Lyberopoulos, D. Tsirkas, and E. Sykas, "Mobility modeling in third-generation mobile telecommunications systems," *Personal Communications, IEEE*, vol. 4, no. 4, pp. 41–56, Aug. 1997.
- [66] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless Communications and Mobile Computing (WCMC), John Wiley & Sons, Ltd.*, vol. 2, pp. 483–502, 2002.
- [67] J. Broch, D. A. Maltz, D. B. Johnson, Y.-C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*, ser. MobiCom '98, 1998, pp. 85–97.
- [68] C. Bettstetter, G. Resta, and P. Santi, "The node distribution of the random waypoint mobility model for wireless ad hoc networks," *Mobile Computing, IEEE Transactions on*, vol. 2, no. 3, pp. 257–269, July-Sept. 2003.
- [69] D. Giancristofaro, "Correlation model for shadow fading in mobile radio channels," *Electronics Letters*, vol. 32, no. 11, pp. 958–959, May 1996.
- [70] X. Cai and G. B. Giannakis, "A two-dimensional channel simulation model for shadowing processes," *IEEE Trans. Vehicular Technology*, vol. 52, no. 6, pp. 1558–1567, Nov. 2003.
- [71] S. Szyszkowicz, H. Yanikomeroglu, and J. Thompson, "On the feasibility of wireless shadowing correlation models," *Vehicular Technology, IEEE Transactions on*, vol. 59, no. 9, pp. 4222–4236, Nov. 2010.
- [72] N. Patwari and P. Agrawal, "Effects of correlated shadowing: Connectivity, localization, and RF tomography," in *Information Processing in Sensor Networks, 2008. IPSN '08. International Conference on*, Apr. 2008, pp. 82–93.
- [73] M. Di Renzo, F. Graziosi, and F. Santucci, "Cooperative spectrum sensing in cognitive radio networks over correlated log-normal shadowing," in *Vehicular Technology Conference, 2009. VTC Spring 2009. IEEE 69th*, Apr. 2009, pp. 1–5.
- [74] B. Kasiri and J. Cai, "Effects of correlated shadowing on soft decision fusion in cooperative spectrum sensing," in *INFOCOM IEEE Conference on Computer Communications Workshops , 2010*, Mar. 2010, pp. 1–6.

- [75] A. H. Al-Dhalaan and I. Lambadaris, “Wireless sensor network localization with spatially correlated shadowing,” in *Communications (ICC), 2010 IEEE International Conference on*, May 2010, pp. 1–6.
- [76] J. H. Lee and R. Buehrer, “Location estimation using differential RSS with spatially correlated shadowing,” in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, Dec. 2009, pp. 1–6.

Curriculum Vitae

Dr. Jaime Almodovar was born in San German, Puerto Rico in 1976. He received a Catholic educational formation, attending Saint Ignatius of Loyola Jesuit school in San Juan, Puerto Rico for his Middle and High School education. In 1999, he received a BSEE from the University of Puerto Rico, Mayagüez Campus. In 2002, he received a MSEE from Johns Hopkins University, Baltimore, MD. Dr. Almodovar has worked for various organizations including the National Aeronautics and Space Administration (NASA), Northrup Grumman Corporation, and The Aerospace Corporation. Most of his work has concentrated in the areas of digital communications and digital signal processing.