The Design and Modeling of Ultra-Wideband Position-Location Networks

Swaroop Venkatesh

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

Committee Members:
Dr. R. M. Buehrer, Chair
Dr. J. H. Reed
Dr. W. A. Davis
Dr. S. F. Midkiff
Dr. B. M. Fraticelli

February 20, 2007
Blacksburg, Virginia

Keywords: Ultra-Wideband, Communication Networks, Position-Location

Design and Modeling of Ultra-Wideband Position-Location Networks

Swaroop Venkatesh

(ABSTRACT)

Impulse-based Ultrawideband (UWB) is a form of signaling which uses streams of pulses of very short duration, typically on the order of a nanosecond. Impulse-based UWB systems possess the ability to fuse accurate position-location with low-data rate communication, and provide covertness for tactical applications and robustness in dense multipath propagation environments. These features can be leveraged in the design wireless ad hoc position-location networks (PoLoNets) for accurate location tracking and monitoring where GPS is not available, especially indoors. Location information is sequentially propagated through a network of reference nodes in order to create a framework for the tracking of mobile nodes, as well as a multi-hop message-passing infrastructure between mobile nodes and control nodes located outside the area of deployment. The applications of such networks include the location and command-and-control of fire-fighters in emergency scenarios, the location of military personnel deployed in urban or indoor environments, and the guidance of robots through large multi-room indoor environments.

The main objective of this dissertation is to derive design principles, techniques and analytical models for UWB PoLoNets that are useful in the development of practical solutions. Some of the fundamental obstacles to obtaining accurate location information in indoor environments are non-line-of-sight (NLOS) signal propagation, limited connectivity between nodes, and the propagation of localization inaccuracies when using sequential estimation approaches in ad hoc scenarios. Several techniques and algorithms that mitigate these effects, thereby allowing the design of PoLoNets with requisite localization accuracy, are presented. Although these techniques are developed from the perspective of a UWB physical layer, the majority are applicable to generic PoLoNets.

Dedication

This work is dedicated to my parents, Amrita, Radha and Ravi.

Acknowledgments

I would like to thank my advisor, Dr. Mike Buehrer, for his vision, guidance and commitment. Throughout my doctoral work, I have learned a great deal about the approach to systematic thought and expression in particular, and about organization and management in general.

I would like to pay tribute to my parents and family, and Amrita Pati, for believing in me, and for their constant encouragement and support.

I extend many thanks to my colleagues at the Mobile and Portable Radio Research Group (MPRG), Virginia Tech, especially Chris Anderson, Jihad Ibrahim, Rekha Menon, Tao Jia, Jeong-Heon Lee, Youping Zhao, Matt Roney, Nathan Harter, John Keaveny, Qiao Chen, Vivek Bharadwaj, Brian Donlan, and Ramesh Palat. I would particularly like to thank Tao Jia and Umair Ahsan for their help in conducting measurements.

I would further like to thank my friends Rakesh Pathak, Arvind Sharma, Satyaprakash Karri, and Srinath Devulapalli for all the (mindless) discussions we had in the belief that our imaginations would be stimulated!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This dissertation deals with the design and modeling of sensor networks used exclusively (or primarily) for the determination of node locations. We call such networks position-location networks (PoLoNets). These networks are assumed (for reasons to be discussed shortly) to use a physical layer based on Ultra-wideband (UWB) signals. Before an explicit thesis statement can be provided, a few preliminaries warrant examination. Section 1.1 discusses the concept, classification and applications of PoLoNets. The applicability of the UWB physical layer to PoLoNets is scrutinized in Section 1.2. The goals of this dissertation are detailed in Section 1.3, and its organization and original contributions are delineated in Section 1.4.

1.1 Position-Location Networks

Position-location has historically been a desired feature in many commercial and military applications. More recently, position location has been an active area of research in many areas including cellular E-911 [3], [4], [5], sensor networks [6], ad hoc networks [7], robotics [8] and ubiquitous computing. The envisioned applications for ad hoc wireless networks often depend on the automatic and accurate location of deployed terminals or nodes, and as a result, there is a rapidly increasing demand in location-based functionality. In sensor networks, particularly for environmental applications [9] such as water quality monitoring, precision agriculture, and indoor air quality monitoring, the available sensing data may be rendered useless by the absence of accurate sensor location estimates. The availability of accurate location estimates of nodes in wireless ad hoc networks can help reduce configuration requirements and device cost in addition to enhancing performance in communication [10] and routing [11]. Additionally, accurate localization enables applications such as inventory management, intrusion detection [12], and traffic monitoring.

We broadly define a generic PoLoNet as a network that has two salient features: (a) the

ability to obtain accurate estimates of the locations of nodes within the area of deployment of the network in the presence of node mobility, and (b) the ability to route these location estimates and other messages to different parts of the network through low data-rate multi-hop communication between nodes of the network. Applications envisioned for PoLoNets include inventory control, home automation, safety networks, tracking personal items, personnel monitoring, command and control in emergency situations, the guidance of robots in remote locations, and many others.

The shift in emphasis on location-estimation in PoLoNets vis-a-vis traditional sensor and mobile ad hoc networks is evident. In sensor and ad hoc networks, location estimation is typically used to enhance the performance of existing protocols, algorithms and techniques, but is not the primary objective of such networks. For instance, the performance of the medium access control (MAC) layer and routing algorithms can be improved with the use of location information, but the data being transported across the network may be of a different nature.

In general, one can view a PoLoNet as a specific type of sensor network where the physical parameters being sensed are the locations of the nodes. However, unlike sensor networks, the expected lifetime of such networks may be limited (e.g., position location in emergency scenarios) and thus, energy efficiency is not the primary metric of interest. On the other hand, PoLoNets differ considerably from typical mobile ad hoc networks where large quantities of data may have to be transported across the network with a certain Quality-of-Service (QOS) while maximizing throughput and/or minimizing latency. In contrast, in PoLoNets, (i) while energy efficiency may be one metric of interest, in a majority of cases, localization accuracy, robustness and scalability, and reliability of communication may take priority over energy efficiency, and (ii) brief messages are assumed to be exchanged between the nodes of the network at low data rates.

The aforementioned differences between PoLoNets and conventional networks motivate the study of the PoLoNet design problem in general, and the differences between the design of PoLoNets and conventional networks in particular. To further emphasize these differences, we describe a sample application for PoLoNets in detail.

Sample Target Application of PoLoNets In outdoor environments, accurate, reliable position information can be obtained via GPS. However, there are many situations where the use of GPS-based location information is either unreliable (e.g., indoor scenarios), or impractical (e.g., where GPS receivers are too bulky or expensive), requiring the development of other solutions. Consider the command and control of a firefighter operation where multiple personnel are deployed inside a building or an *area of interest*[1]. For safety and efficiency purposes, it would be extremely helpful for a command center outside the building to be established for not only communication but also position tracking as shown in Figure

[1]We focus mainly on two-dimensional location-estimation and tracking.

1.1. In such a case, we require an *ad hoc position location and communication network* that is independent of GPS and is not reliant on pre-existing infrastructure.

The goals of a PoLoNet applicable in this scenario would be to: (a) equip each firefighter with awareness of his own location, (b) equip command-and-control with knowledge of each firefighter's location, and (c) allow the exchange of short messages between firefighters and command-and-control.

Therefore, a PoLoNet used for such applications is required to provide location-awareness within the area of interest, in addition to serving as a low data-rate communication network. The design of PoLoNets for such applications represents a particularly challenging problem due to (a) the necessity for extremely accurate node location estimates in dense propagation environments, (b) the lack of pre-existing localization infrastructure, and (c) constraints on the latency of such networks. In the following section, we classify PoLoNets that can be used for the application described above, based on different criteria.

1.1.1 Classification of PoLoNets

PoLoNets that can be used in the application described above can be classified on the basis of several criteria:

- **Physical Layer:** Traditional ranging (and position location) applications have relied on optical (laser), ultrasound, or narrowband RF physical layers. It is well known that optical and ultrasound ranging have limited range in harsh environments and may fail completely when the line-of-sight (LOS) is blocked. Additionally, narrowband RF solutions are adversely affected in dense multipath due to severe multipath fading. As will be discussed in Section 1.2, UWB is an excellent physical layer solution for indoor PoLoNets because of its usefulness in harsh multipath environments, material penetration capabilities, its ability to fuse *accurate position-location* with *low-data rate communication* and its covertness for tactical applications.

- **Measured information:** Information of several kinds can be used to estimate node locations. The most important categories are: (i) Time-of-Arrival (TOA) based range estimates, (ii) Time-Difference-of-Arrival (TDOA) based range estimates, (iii) Received Signal Strength (RSS) based range estimates, (iv) Angle-of-Arrival (AOA) information, and (v) Connectivity based range information, to name a few. UWB signals can be used for accurate TOA or TDOA-based range estimation. RSS-based range estimation is typically used in narrowband sensor networks.

- **Infrastructure-based Vs. Ad hoc PoLoNets:** An *infrastructure-based* PoLoNet is a network of nodes whose locations are precisely known, being deployed within the area of interest in advance. Such a network can directly assist in the localization of mobile nodes within the area of interest through the multi-lateration of available range

information. In the most general sense, GPS can be thought of as an infrastructure-based PoLoNet. In contrast, nodes whose locations are unknown *a priori* can be deployed at the time of use in an *ad hoc* fashion. These nodes subsequently estimate their locations using range information from a small number of localized[2] nodes or "anchors", typically located outside the area of interest. Ad hoc PoLoNets have several practical applications, and are discussed in detail in the following subsection.

- **Synchronous Vs. Asynchronous:** When all nodes in a PoLoNet share a common clock, they are said to be synchronous; if each node possesses a unique clock, they are said to be asynchronous. Synchronization between nodes can allow for more efficient and accurate range estimation between nodes, as we shall see subsequently.

- **Centralized Vs. Distributed Location Estimation:** Due to constraints on the hardware complexity of nodes, or limited connectivity to other localized nodes, it may not be possible for each node to estimate its location based solely on locally available information. In such cases, range and location information can be relayed to a centralized "location-solver" that jointly estimates the locations of several unlocalized nodes. In *distributed location-estimation*, unlocalized nodes estimate their locations independently, based only on *local* range and location information.

- **Quasi-static Vs. Dynamic Scenarios:** In many applications, the locations of the nodes being tracked may be static or changing very slowly. The rate of change of these locations determines the rate at which location estimates need to be updated. This implies that the estimation of the locations of nodes in a quasi-static scenario needs to be performed less frequently than in a dynamic scenario with high node mobility.

In the following section, we describe an ad hoc PoLoNet architecture and the corresponding mechanisms that can be applied to practical scenarios, such as the fire-fighter tracking and communication problem discussed in Section 1.1.

1.1.2 Ad Hoc UWB PoLoNets

The availability of an infrastructure-based PoLoNet allows the direct tracking of mobile nodes in the area of interest. However, in most cases, it may not be possible to have an infrastructure of stationary location-aware nodes within the area of interest *a priori*. Further, due to (i) the nature of the indoor propagation environment and (ii) the transmit power restrictions, we would require the PoLoNet to comprise a *multi-hop* network of localized nodes in order to achieve network-wide communication. A network architecture that incorporates

[2] A node is said to be "location-aware" if its location is known *a priori* and "localized" if its location is known or can be estimated based on available range information. A node whose location is unknown is said to be "unlocalized".

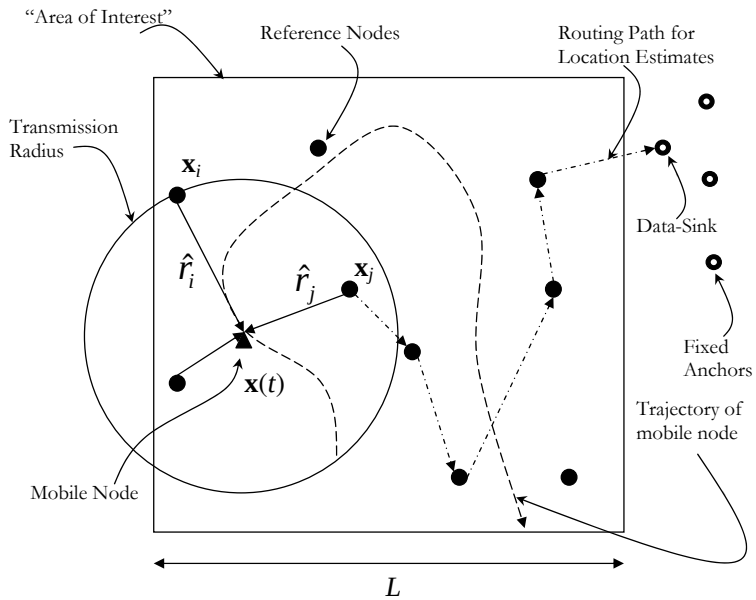


Figure 1.1: Ad hoc PoLoNet architecture: the data-sink shown represents the command and control station. Reference nodes, whose locations are known *a priori* initially estimate their locations using local range information. These localized reference nodes then assist mobile nodes in estimating their own locations by providing range and location information, while serving as a multi-hop message passing architecture between mobile nodes and the data sink.

these features is illustrated in Figure 1.1. The network consists of a small number of location-aware *fixed anchors* located outside the area of interest. The locations of fixed anchors may be available via GPS, or by setting up a local coordinate system. Starting from the setup instant, the network evolves in two phases:

Phase 1 Nodes called *reference nodes* or *propagated anchors*, whose locations are unknown *a priori*, are deployed[3] in the area of interest. The deployed reference nodes, depending on the available connectivity to other nodes, gather range and location estimates, which can then be used to estimate node locations in a distributed manner. We assume that reference nodes triangulate their locations in a distributed manner using range estimates from fixed anchors or other reference nodes whose locations have already been estimated. Reference nodes that estimate their locations provide range estimates to other unlocalized reference or mobile nodes. Each "layer" of location-aware reference nodes serves as a source of range information for the subsequent layer, thereby sequentially propagating location-awareness, even in the absence of direct connectivity with fixed anchor nodes. It should be evident that

[3]Deployment options are not considered here but reference nodes could either be pre-existing, deployed manually as in a fire-fighter scenario, via robots, dispersed via UAV, or launched into the area of interest.

this phase results in the formation of a network of localized reference nodes[4] that can be used to track the locations of mobile nodes.

Phase 2 After reference nodes estimate their own locations by ranging to one another or to fixed anchors, the second phase of the network involves assisting any mobile node that enters the area of interest by providing a framework to estimate its location. Mobile nodes, depending on their location and available connectivity, communicate with a subset of fixed anchors and/or localized reference nodes in order to obtain range information. These range estimates are then used to triangulate their locations. In this phase, the reference nodes additionally provide a multi-hop communication network to relay the mobiles' location information to, and short messages from, a data-sink (command-and-control).

In this manner, through the network of reference nodes, (i) location-awareness is propagated from the fixed anchors located outside the area of interest to the mobile nodes within the area of interest, (ii) mobile node location information is passed from the mobile nodes to the data-sink, and (iii) messages are relayed between the data sink and mobile nodes. An illustration of the two network phases is shown in Figure 1.2.

In the following section, we briefly discuss impulse-based UWB technology, a physical layer whose features are ideally suited to the design of PoLoNets.

1.2 Ultra Wideband Technology

The U.S. Federal Communications Commission (FCC) has defined an Ultra-wideband (UWB) device to be any intentional radiator of radio frequency (RF) energy, which has a 10 dB bandwidth of 25 percent of the strongest frequency within that 10 dB bandwidth, or a 10 dB bandwidth equal to or greater than 500 MHz. Assuming that the upper and lower limits of the 10 dB bandwidth of the signal are f_H and f_L respectively, and that the strongest frequency lies at the center of the 10 dB bandwidth, the definition of a UWB signal can be expressed as:

$$2\left(\frac{f_H - f_L}{f_H + f_L}\right) \geq 0.25 \text{ or } (f_H - f_L) \geq 500 \text{ MHz}. \tag{1.1}$$

UWB signals have been used in the military for radar applications since the 70's. In February 2002, the FCC legalized the use of UWB by releasing a set of spectral "masks", in the portion of the spectrum from several hundred MHz to approximately 10 GHz, which stipulated the emission level and frequency of operation for imaging, radar, and communication purposes. This sparked a huge interest in UWB technology.

[4]This can be thought of as the formation of a "noisy" infrastructure-based PoLoNet.

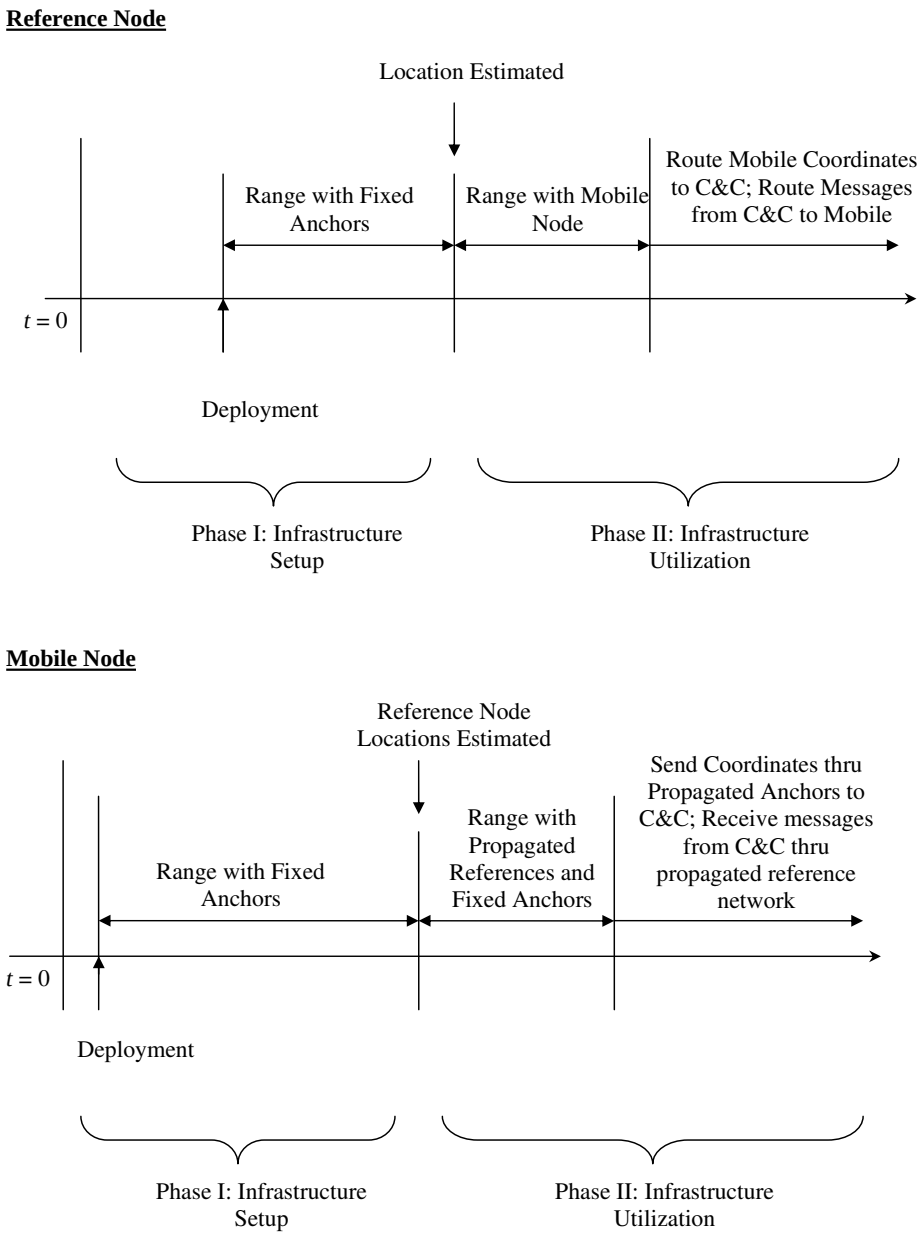


Figure 1.2: Roles of Nodes in different PoLoNet phases: In phase 1, reference nodes attempt to estimate their locations to form an infrastructure-based PoLoNet that can assist mobile nodes in estimating their own locations in phase 2. In phase 2, the reference nodes additionally serve as a multi-hop communication network between the mobiles and the data sink.

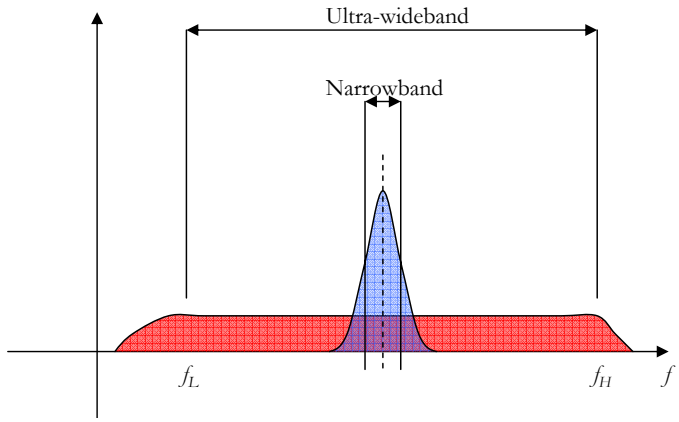
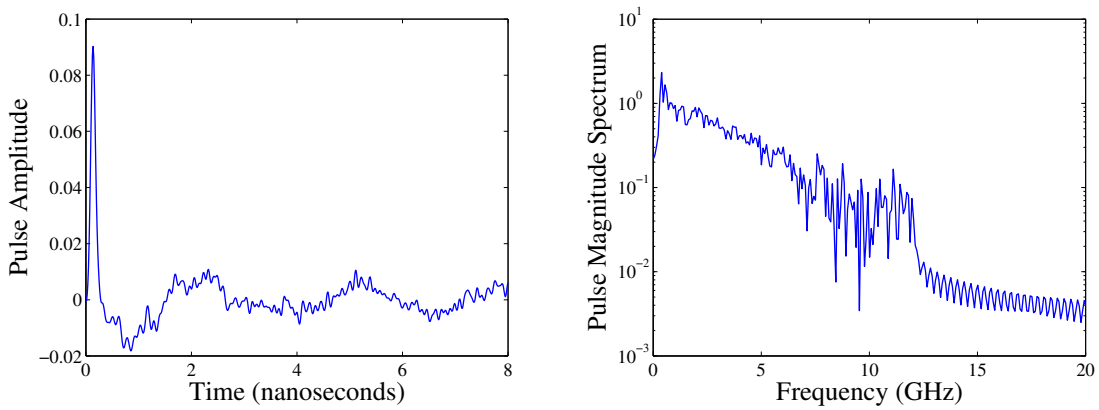


Figure 1.3: Ultra-wideband versus Narrowband signals

Impulse-based UWB or Impulse Radio [13] is a form of UWB signaling which uses streams of pulses of very short duration, typically on the order of a nanosecond, thereby spreading the power spectral density of the radio signal over several gigahertz as illustrated in Figure 1.3. Sample UWB signals in the time and frequency domains are shown in Figure 1.4. Since UWB systems have to operate in the highly populated frequency range below a few gigahertz, UWB radios must not only contend with a variety of interfering signals, but must also ensure that they do not interfere with narrowband radio systems operating in dedicated bands by satisfying the restrictions on the transmit power spectral density imposed by the FCC. These constraints on the transmit power spectral density necessitate the reduction of data rates to improve the coverage radius of transmissions. This directly leads to the use of direct-sequence (DS) or time-hopping (TH) [13] spread-spectrum techniques applied to the low-duty cycle UWB pulse trains, with data modulation accomplished by pulse-amplitude modulation (PAM) or pulse-position modulation (PPM) with several pulses per data symbol. Multi-band OFDM (MB-OFDM) is another form of UWB signaling where the entire bandwidth is divided into sub-bands of approximately 500 MHz, and orthogonal frequency division multiplexing is used within each sub-band in order to provide high data rates over short distances.

Impulse-based UWB has some unique advantages over traditional narrowband systems [1] that include:

- The received pulses are relatively immune to the multipath fading that afflicts narrowband signals, as seen in Figure 1.5. The robustness to multipath fading is due to the fine temporal resolution provided by UWB signals, which allows individual multipath components to be resolved. Additionally, UWB signals have significant material penetration capabilities (particularly in the lower bands approved by the FCC), which is desirable in indoor networks.



(a) Time Domain

(b) Frequency Domain

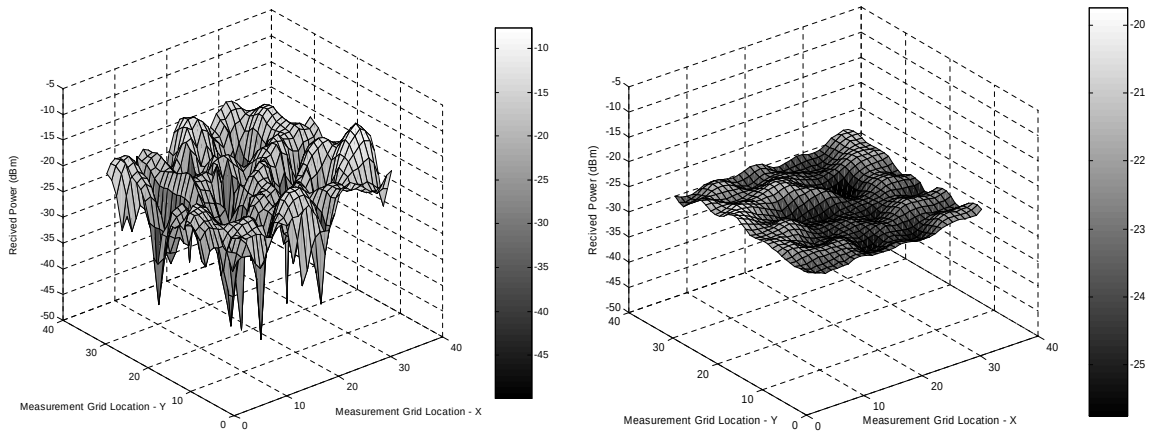
Figure 1.4: i
n the time and frequency domains.]Ultra-wideband pulses [1] in the time and frequency domains.

- Due to the short duration of the pulses, UWB can be used for precise ranging. This feature can be leveraged in the positioning and tracking of devices, and offers the potential to fuse communications, positioning and sensing functionalities.
- Impulse-based UWB has the potential to support low to medium data rates while appearing noise-like to other RF technologies, and offers inherent data security due to covertness.

In the following section, we discuss the challenges of designing PoLoNets based on the UWB physical layer, and the goals of this dissertation in that regard.

1.3 Macroscopic Design Framework for UWB Ad Hoc PoLoNets

In this dissertation, the research challenge being addressed is the modeling and design of PoLoNets using UWB signals in indoor environments without the need for fixed infrastructure. Specifically, we focus on the design of *UWB ad hoc TOA-based* PoLoNets from the perspective of localization accuracy. This dissertation then identifies the fundamental limitations to the localization accuracy achievable through indoor UWB PoLoNets and provides initial solutions that address these limitations. For the target application described in Section 1.1, the localization accuracy of mobile nodes that can be guaranteed over the entire



(a) Narrow-band Signal

(b) Ultra-wideband Signal

Figure 1.5: .
]Spatial fading of Narrow-band and Ultra-wideband signals over a given area [1].

area of interest is a key metric of interest[5], and is likely to be a crucial parameter in PoLoNet design.

The goal of this dissertation is to broadly answer three main questions regarding the design of indoor UWB PoLoNets from the perspective of localization accuracy: *(a) Is it possible to design PoLoNets using a UWB physical layer that satisfy the practical design constraints on localization accuracy? (b) What are the main factors that assist/impede the achievement of the target localization accuracy? and (c) Can techniques that mitigate these impediments be developed for UWB PoLoNets in order to ensure the desired localization accuracy?*

The potential PoLoNet design issues that relate to (i) the requirement of accurate, robust and rapid location estimation, and (ii) the use of the UWB physical layer, that are addressed in this dissertation are briefly summarized below:

- **Range Estimation:** The accuracy of range estimation possible through the use of UWB signals in different propagation scenarios needs to be quantified.
- **Location Estimation:** Given the achievable range measurement accuracy using UWB signals, the achievable localization accuracy needs to be quantified. The factors that impact localization accuracy need to be analyzed, and their effect quantified. Further, the use of "practical" low-complexity location estimators that provide sufficiently

[5]For instance, in a fire-fighter [14] position-tracking system, the knowledge of whether a firefighter is on one side of a door or the other, could be critical. Thus, the target localization accuracy (in terms of the area of uncertainty of location estimates) for such applications is likely to be on the order to 1 meter^2.

accurate location estimates needs to be studied.

- **Robustness in NLOS propagation environments:** In indoor PoLoNets, the direct use of range information in NLOS propagation scenarios can result in considerable degradation in localization accuracy. This necessitates the characterization of the impact of NLOS signal propagation on localization accuracy, and the development of NLOS mitigation algorithms.

- **Propagation of Error:** In an ad hoc PoLoNet with *sequential* distributed location estimation discussed in Section 1.1.2, the propagation of localization error is a key concern. All reference nodes may not have connectivity with fixed anchors, and therefore, need to estimate their locations based on other localized reference nodes. As these estimates are likely to be noisy, as the distance to the fixed anchors increases, these errors can accumulate. This results in poor localization accuracy in regions of the area of interest far away from fixed anchors. The extent of the propagation of localization error needs to be quantified, and the methods that mitigate these effects need to be developed.

- **Multiple-Access Design:** Multiple-access schemes in sensor and mobile ad hoc networks are typically designed from the perspectives of data throughput, latency or energy efficiency. As localization accuracy is the main metric of interest in PoLoNets, the connection between multiple-access protocols and localization accuracy needs to be scrutinized. Further, the functional relationship between various network parameters and localization accuracy warrants further investigation. Multiple-access protocols that are optimized with respect to (a) localization accuracy, and (b) the UWB physical layer need to be investigated.

- **Power Control Algorithms:** As a mobile node moves through the area of interest, the connectivity with localized reference nodes can fluctuate considerably. Further, if the probability of reference node failure is significant (e.g., fire-fighter tracking network), controlling the transmit power to ensure the minimum required connectivity becomes essential. Therefore, the effect of power-control on localization accuracy needs to be quantified, and efficient power-control algorithms need to be developed.

In this work, we analyze the connections between localization accuracy and different PoLoNet design elements with respect to the issues discussed above, as shown in Figure 1.6. This allows for a unified treatment of the design problem for UWB PoLoNets from the perspective of localization accuracy. Further, we are interested in developing techniques, algorithms and protocols that can enhance the performance of the design elements in Figure 1.6, as well as the prediction of their performance under diverse conditions. Such a unified design framework allows for a "localization-budget"—analogous to a link-budget in communication system design—that provides an evaluation of the design requirements in order to obtain a target localization accuracy. Finally, the insight into the factors that affect localization

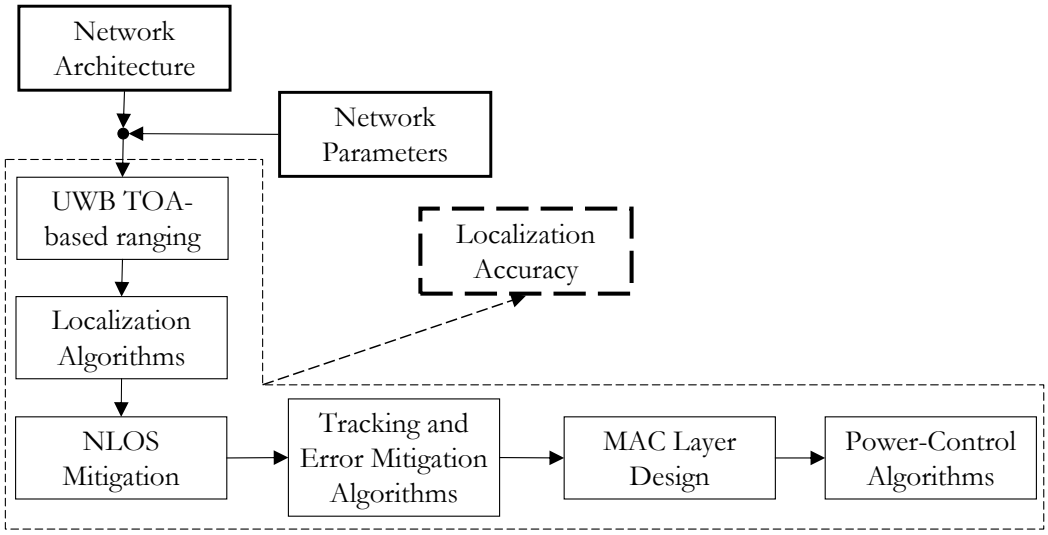


Figure 1.6: Unified design of UWB PoLoNets with respect to localization accuracy

accuracy allows for the active selection of algorithms, thereby controlling the trade-offs between performance and complexity. It is important to point out that although the emphasis in this dissertation is on the design of UWB PoLoNets, a vast majority of the insights gained and techniques developed are valid for any networks determining node position in general.

The following section details the contents of this dissertation and delineates previous and original work.

1.4 Organization and Original Contributions

The problem of modeling and design for UWB PoLoNets is fairly recent [15], and the literature on the subject is scarce. However, the relevance of existing literature on the UWB physical layer, and literature on localization in cellular, sensor, and other networks needs to be addressed. In order to clearly distinguish previous (background) work from original work, in the following we highlight the main original contributions in each chapter, and the previous work that is included in each chapter for the sake of completeness and consistency. This dissertation is organized as follows:

In Chapter 2, we provide a characterization of the performance of the efficient estimator for synchronized TOA-based ranging in AWGN and in the presence of multipath, through the derivation of the Cramer-Rao lower bound for the variance of an unbiased estimator. We then describe a two-way packet handshake for distributed ranging in the absence of synchronization between nodes. A practical non-efficient estimator based on energy thresholding of

the received multipath profile is then analyzed in detail. We then compare the "ideal" (efficient) range estimator with practical estimators in measured multipath channels, to derive a model for range estimates that is used in the rest of the work.

Previous Work in Chapter 2: CRLB for synchronous TOA estimation [16], and for UWB signals in multipath [17]. Two-way packet handshake mechanism for asynchronous ranging [18].

Original Contributions in Chapter 2:

- Analysis of a non-efficient estimator (based on energy thresholding) for range using TOA information.
- Comparison of the efficient and non-efficient estimates in multipath channels through measurements.

Chapter 3 uses the model for the accuracy of range information derived in Chapter 2 to arrive at bounds on the performance of location estimation. A systematic formal study of the properties of the CRLB on the localization error in terms of the variance of range estimates, the impact of node geometry and the number of range estimates and other parameters, is presented. The insights into the properties of the CRLB are used in establishing the connection between localization accuracy and multiple access schemes and power-control algorithms. We then compare the performance of practical estimators based on a Least-squares (LS) approach with the CRLB in order to (i) verify that similar trends were observed in terms of the properties, and (ii) quantify the difference between the CRLB and practical estimators.

Previous Work in Chapter 3: CRLB for location estimation with unbiased Gaussian range estimates [9], [19], the Least-squares estimator [20], Iterative refinement using Levenberg-Marquhardt iteration [21].

Original Contribution in Chapter 3:

- The systematic study of the properties of the CRLB provides insights into the factors than impact localization accuracy. These trends were shown to be followed by practical estimators, and are used in establishing the connection between localization accuracy and multiple access schemes and power-control algorithms [22], [23].

In Chapter 4, the impact of NLOS propagation on localization accuracy is examined. Addressing the impact of NLOS propagation requires partitioning the problem into two subproblems: NLOS identification and NLOS mitigation. The former deals with the problem of distinguishing between LOS and NLOS range estimates, whereas the latter typically deals with the reduction of the adverse impact of NLOS range information on the accuracy of

location-estimates, assuming the NLOS range estimates have been identified. We investigate the performance of statistical hypothesis-testing of received signal parameters for NLOS identification. We propose and evaluate a novel NLOS identification method based on delay spread statistics that can be extracted from the received multipath profile.

Assuming NLOS range estimates can be identified, the CRLB with a mixture of LOS and NLOS range estimates is interpreted. We then explore the impact of the NLOS range estimates on the localization accuracy achieved by practical estimators. We then propose a novel linear programming approach to NLOS mitigation that takes advantage of the features of UWB signals, and provide an exhaustive evaluation of its performance in different scenarios via simulations.

Previous Work in Chapter 4: The impact of NLOS range estimates on localization accuracy has been considered in the context of cellular communications [4], [24]. The CRLB with biased Gaussian NLOS range estimates has been derived in [25].

Original Contributions in Chapter 4:

- The development and analysis of a novel NLOS identification method based on received signal statistics [26].
- The development and analysis of a novel NLOS mitigation technique based on linear programming [27], [28].

Chapter 5 presents a systematic study of the bounds on the performance of different estimation approaches (fully-distributed, sequential-distributed and centralized), followed by the comparison of practical implementations of these approaches. Insights into the nature of the propagation of localization error, and its impact on the design of ad hoc PoLoNets are provided. These insights allow us to enumerate various means of limiting the propagation of localization error. The propagation of localization error in NLOS environments has not been studied previously. A novel method of mitigating the propagation of localization error, that utilizes range estimates over multiple hops while incorporating NLOS range estimates, is proposed.

Previous Work included in Chapter 5: The CRLB for centralized estimation was derived in [19].

Original Contributions in Chapter 5:

- Insights into the nature of the propagation of localization error, and its impact on the design of ad hoc PoLoNets are provided. Based on these insights, various means of limiting the propagation of localization error are enumerated.
- A novel method that mitigates the propagation of localization error in NLOS environments by utilizing range estimates over multiple hops is proposed and analyzed.

Chapter 6 presents full-fledged indoor measurement results that demonstrate that UWB PoLoNets have the potential to achieve practical target localization accuracies. The efficacy of the proposed NLOS identification, NLOS mitigation, propagation of error mitigation algorithms, and the use of Kalman smoothing techniques is demonstrated through measurements.

Previous Work in Chapter 6: Kalman smoothing for location estimates in cellular networks has been studied in [29].

Original Contributions in Chapter 6:

- A complete evaluation of the performance of an indoor UWB PoLoNet based on measurements.
- Validation of the proposed NLOS identification, NLOS mitigation, and propagation of error mitigation algorithms through measurement results [26].

Chapter 7 discusses the problem of MAC protocol design specifically for UWB PoLoNets, and specifies the required characteristics of a MAC protocol for UWB PoLoNets. We then use results derived in Chapter 3 to establish a connection between localization accuracy and the selection of MAC protocols. Based on this connection, a spread-spectrum multiple access scheme for UWB PoLoNets is proposed, which is shown to outperform the traditional Carrier-Sense Multiple Access (CSMA) protocols in terms of the convergence of node location estimates to their true values. Analytical models are derived in order to obtain bounds on the performance of the proposed protocol.

Previous Work in Chapter 7: The Common-Transmitter protocol was proposed for Spread-Spectrum systems in [30].

Original Contributions in Chapter 7:

- Establishing connection between MAC protocol design and localization accuracy through the properties of bounds on location estimation.
- The specification of a modified spread-spectrum MAC protocol [23] for distributed ranging in UWB PoLoNets, and the analytical modeling of the proposed scheme followed by comparison the traditional CSMA protocol via simulations.

Chapter 8 discusses the dependence of localization accuracy on network parameters such as node densities, mobile speed and other multiple-access parameters. We consider both synchronous and asynchronous networks, for the case of ground-deployed nodes. The impact of network parameters on the localization accuracy of a node as a function of time is investigated via analysis, and compared with simulation results, in order to provide insights

into the relationship between localization accuracy and network parameters.

Previous Work in Chapter 8: The distribution of the received power from a Poisson field of interferers has been derived in [31].

Original Contributions in Chapter 8:

- A framework for the analytical modeling of the relationship between multiple-access system parameters and the resulting average localization accuracy in synchronous and asynchronous PoLoNets.
- Insights into the impact of the density of localized nodes and speed of mobile nodes on average localization accuracy based on analysis, and verification via simulation results.

Chapter 9 discusses the need for, and implementation of, power-control algorithms in UWB PoLoNets. We examine the variation of localization accuracy due to node mobility, and the impact of power-control on the localization accuracy. A novel method of quantifying and estimating the "quality" of location-estimates is discussed, and two power-control schemes based on the defined quality metric are detailed. The stochastic convergence of one of proposed power-control approaches is demonstrated analytically in single-user scenarios, and evaluated via simulations for the multi-user case. A comparison of these power-control approaches with simulated optimal solutions is provided.

Previous Work in Chapter 9: A framework for analyzing the stochastic convergence of power-control algorithms in cellular systems was developed in [32], [33].

Original Contributions in Chapter 9:

- The impact of power-control on localization accuracy is analytically characterized [22].
- A metric for assessing the quality of a node's location estimate is proposed, and an iterative power control scheme is developed based on such a quality metric. The stochastic convergence of this scheme is verified for a single-user scenario.
- A second non-iterative power control scheme is proposed based on the above metric, and the properties of the CRLB, whose performance is verified via simulations [22].

Chapter 10 concludes this dissertation and presents an overview of the ideas and insights gained on the design of UWB PoLoNets.

Chapter 2

TOA-based Ranging using UWB

2.1 Introduction

The process of node localization typically requires the use of distance, connectivity or angle-of-arrival information. Given a sufficient amount of such information, the location of a node can be estimated or constrained. Figure 2.1 illustrates the use of distance, connectivity or angular information from three anchor nodes in the localization of a node. Distance or *range* information can be obtained through the estimation of the time-of-arrival (TOA) or received signal strength (RSS) of signals from anchors. For localization using UWB radios, TOA-based ranging is expected to provide very good accuracy due to the high time resolution (large bandwidth) of UWB signals. Furthermore, TOA-based schemes are likely to be less expensive than the AOA-based schemes, which typically require the provision of multiple antennas, and are less effective in dense scattering environments. While the estimation of the distances based on RSS estimation is computationally efficient, the requirement of extremely accurate path loss models limits the accuracy of RSS-based range estimation in dense multipath environments. While the use of connectivity information can be used to constrain the location of a node, location estimates obtained from connectivity information is very coarse compared to those obtained from the TOA measurements. Due to the inherent suitability and accuracy of TOA-based approaches for UWB systems, we focus our attention on TOA-based ranging.

The goal of this chapter is to characterize the accuracy of TOA-based ranging using UWB signals so as to provide a basis for the characterization of localization accuracy, the subject of chapter 3. We begin by characterizing bounds on the accuracy of *synchronized* TOA-based range estimation, and analyzing the impact of SNR, bandwidth, multipath channels, and NLOS propagation on range estimation. We then describe a ranging mechanism that can be used in the absence of synchronization between nodes. A detailed investigation of a computationally efficient TOA estimator for UWB signals using the received multipath

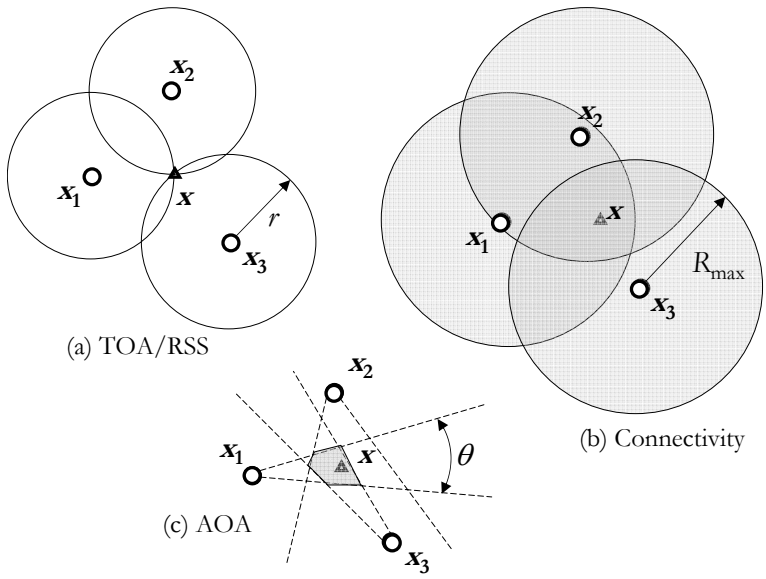


Figure 2.1: Various kinds of information that can be used for localization: (a) TOA/RSS information, (b) Connectivity information, (c) AOA information.

profile is then presented. We then compare the "ideal" (efficient) range estimator with practical estimators in measured multipath channels, to provide a model for range estimates that is used in the rest of the work.

Previous Work in this Chapter: CRLB for synchronous TOA estimation [16], and for UWB signals in multipath [17]. Two-way packet handshake mechanism for asynchronous ranging [18].

Original Contributions in this Chapter: Analysis and Simulation of a non-efficient estimator (based on energy thresholding) for range using TOA information. Evaluation of a practical estimator in UWB multipath channels through measurements.

The organization of this chapter is as follows: in Section 2.2, we provide a characterization of the performance of the efficient range estimator for synchronized TOA-based ranging in AWGN and in the presence of multipath, through the derivation of the Cramer-Rao Lower Bound for the variance of an unbiased range estimator. Section 2.3 describes a two-way packet handshake for distributed ranging in the absence of synchronization between nodes. Section 2.4 presents a simple TOA estimator based on energy thresholding of the received multipath profile, which is analyzed in detail. Finally, we compare the "ideal" (efficient) range estimator with practical estimators in measured multipath channels, to present the rationale a model for range estimates in LOS and NLOS propagation environments that is used in the rest of the work in Section 2.5. This chapter concludes in Section 2.6.

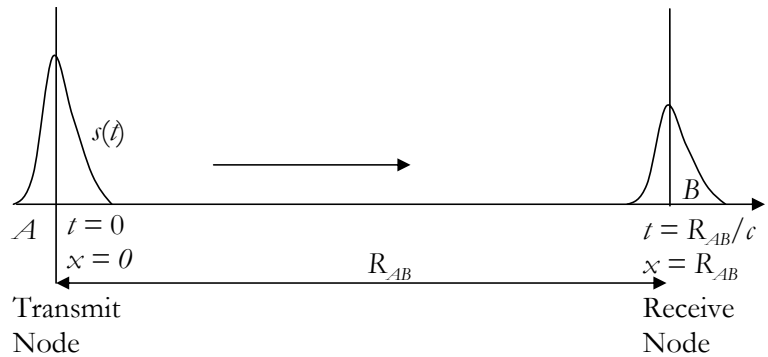


Figure 2.2: Illustration of TOA-based ranging

2.2 Bounds on TOA-based Range Estimation

Detection and estimation problems associated with a signal traveling between nodes have been extensively studied [16] in radar and other applications. TOA-based ranging techniques rely on a measurement of the travel time of a signal transmitted between nodes. If two nodes have a common time-reference, the node receiving the signal can determine the TOA of the incoming signal as shown in Figure 2.2. Suppose the signal $s(t)$ is transmitted by a node A at $t = 0$. This signal is received at node B corrupted by noise, after a propagation delay $\tau_0 = \frac{R_{AB}}{c}$, where c is the speed of light:

$$r_B(t) = s(t - \tau_0) + n(t)$$

where $n(t)$ is assumed to represent AWGN. Assuming we have an estimate $\hat{\tau}_0$ of the TOA τ_0 of the signal at the receiver, then the range can be estimated using

$$r_{AB} = c\hat{\tau}_0 \tag{2.1}$$

It is evident that the accuracy of the range estimate r_{AB} is strongly dependent on the accuracy of the TOA estimate $\hat{\tau}_0$, since the variance of the range estimate is related to the variance of the TOA estimate by $\sigma_r^2 = c^2\sigma_\tau^2$.

The Cramer-Rao lower bound (CRLB) for a parameter that is to be estimated specifies the *minimum variance of an unbiased estimator* [34] of that parameter. The analysis leading to the CRLB does not specify the estimator that achieves the CRLB. However, the CRLB serves as a benchmark for the performance of practical estimators, as the accuracy of an available unbiased estimator can be evaluated based on the difference between its variance and the CRLB. An estimator that attains the CRLB is said to be *efficient*. In the following subsections, we look at bounds on the accuracy of TOA-based range estimation in UWB channels using CRLB analysis.

2.2.1 Range Estimation in AWGN

For a single-path line-of-sight (LOS) additive white Gaussian noise (AWGN) channel, the signal transmitted by node A in Figure 2.2 is assumed to be of the form:

$$s(t) = \sum_{k=1}^{N_m} p(t - kT_f), \tag{2.2}$$

where $p(t)$ is the transmit pulse shape, N_m is the number of pulses, and T_f is the pulse-repetition interval. It is assumed that the pulse width T_w is much smaller than T_f: $T_w \ll T_f$. In general, we can assume the presence of a direct-sequence (DS) or time-hop (TH) code [35], resulting in a transmit signal of the form:

$$s(t) = \sum_{k=1}^{N_m} a_k p(t - kT_f - c_k T_c),$$

where a_k and c_k are the DS and TH "chips" respectively, and T_c is the hop interval. For simplicity, we assume the transmit signal represented by (2.2). The corresponding received signal at node B is then modeled using

$$r_B(t) = \gamma s(t - \tau_0) + n(t) = \gamma \sum_{k=1}^{N_m} p(t - kT_f - \tau_0) + n(t), \tag{2.3}$$

where γ represents the gain of the channel, τ_0 denotes the propagation delay, and $n(t)$ represents zero-mean AWGN with variance σ_n^2.

Assuming a common time reference shared by nodes A and B, and given the above received signal of duration $T_m = N_m T_f$, an estimate $\hat{\tau}_0$ of the propagation delay τ_0 can be computed. For an unbiased estimate $\hat{\tau}_0$ with variance σ_τ^2, the following properties are known [16], [17] to hold:

Theorem 1 (Range Estimation Bound in AWGN). *The CRLB for an unbiased estimate of a time-of-arrival (TOA) of a single multipath component in AWGN is given by*

$$\sigma_\tau^2 \geq \frac{\sigma_n^2}{\gamma^2 \int_{T_m} \dot{s}^2(t)dt}, \tag{2.4}$$

where $\dot{s}(t)$ denotes the time-derivative of the transmitted signal.

From the above theorem, for the transmit signal defined in (2.2), we have:

$$\sigma_\tau^2 \geq \frac{\sigma_n^2}{\gamma^2 N_m \int_{T_f} \dot{p}^2(t)dt} = \frac{\sigma_n^2}{\gamma^2 N_m E_p} \frac{\int_{T_f} p^2(t)dt}{\int_{T_f} \dot{p}^2(t)dt}, \tag{2.5}$$

where E_p is the energy per pulse. For a real pulse $p(t)$ with spectrum $P(f)$, since $T_w \ll T_f$, we can write

$$\frac{\int_{T_f} p^2(t)dt}{\int_{T_f} \dot{p}^2(t)dt} = \frac{\int_{-\infty}^{\infty} |P(f)|^2 df}{\int_{-\infty}^{\infty} 4\pi^2 f^2 |P(f)|^2 df} = \frac{1}{4\pi^2 B_e^2},$$

where B_e is the effective root-mean-squared (RMS) bandwidth of the pulse defined as:

$$B_e \triangleq \sqrt{\frac{\int_{-\infty}^{\infty} f^2 |P(f)|^2 df}{\int_{-\infty}^{\infty} |P(f)|^2 df}}. \tag{2.6}$$

Further, in (2.5), we note that the received signal-to-noise ratio (SNR) ξ is given by:

$$\xi = \frac{\gamma^2 N_m E_p}{\sigma_n^2}.$$

Therefore, we have:

$$\sigma_\tau^2 \geq \frac{1}{4\pi^2 \xi B_e^2}.$$

Since the range estimate r is obtained using $r = c\hat{\tau}$, we obtain the following bound for an unbiased range estimator:

$$\sigma_r^2 \geq \frac{c^2}{4\pi^2 \xi B_e^2}. \tag{2.7}$$

In the absence of multiple-access interference, the received SNR can be modeled as:

$$\xi = K_P P_T R^{-\beta}, \tag{2.8}$$

where P_T is the transmit power, β is the path-loss exponent, and K_P is a constant that subsumes the effect of other physical layer parameters such as N_m, the antenna gains, and the receiver noise floor. From (2.7) and (2.8),

$$\sigma_r^2 \geq \frac{c^2 R^\beta}{4\pi^2 B_e^2 K_P P_T}. \tag{2.9}$$

Corollary. *For the UWB signals defined in this work, the minimum variance of an unbiased range estimator can be written in two equivalent forms:*

$$\begin{aligned} \sigma_r^2 &\geq \frac{K'_R}{\xi} = \frac{K'_R R^\beta}{K_P P_T}, \text{ where } K'_R = \frac{c^2}{4\pi^2 B_e^2}, \\ \sigma_r^2 &\geq K'_E R^\beta \text{ where } K'_E = \frac{K'_R}{K_P P_T}. \end{aligned} \tag{2.10}$$

From the above results, we see that the accuracy of range estimation (i) improves with increasing SNR, and (ii) improves with decreasing pulse-width (or increasing effective RMS bandwidth B_e). Further, as the SNR at the receiver decreases with increasing R, we see that the accuracy of the range estimate degrades as the distance between nodes increases. The large bandwidth of UWB signals allows extremely accurate TOA-based range estimates over short distances, as illustrated in the following example.

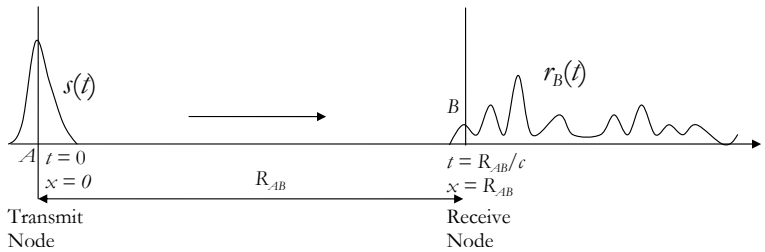


Figure 2.3: Illustration of TOA-based ranging in multipath

Example: TOA-based Ranging with UWB in AWGN With a Gaussian UWB pulse of width $T_w = 0.5$ nanoseconds, the effective RMS bandwidth B_e can be estimated using (2.6) as 1.35 GHz. Under FCC limits, the transmit power spectral density is -41.3 dBm/MHz between 3.1 and 10.6 GHz. Based on these numbers, the maximum transmit power over this band is approximately $P_T = 0.5$ mW. Assuming the received SNR ξ at $R = 1$ meter is 20 dB, we obtain $K_P = 2 \times 10^6$. Using these values, $K_E \approx 1.25 \times 10^{-6}$. At this value of K_E and $\beta = 2$, the minimum standard deviation of an unbiased range estimate predicted by (2.10) at a distance $R = 10$ meters under LOS conditions is 1.1 centimeters.

The Maximum Likelihood (ML) estimate [34] of the TOA of a single path in AWGN can be shown [36] to be reduced to an estimate based on the output of a matched filter or correlation receiver. The TOA τ_0 is estimated by correlating the received signal $r_B(t)$ with the transmitted pulse $p(t)$ and extracting the time-instant at which the correlation is maximized. It is also well known in estimation theory [34] that the ML estimate asymptotically achieves the CRLB as the amount of available data, and therefore the effective SNR, increases.

2.2.2 Range Estimation in Multipath

In the previous development, the CRLB for range estimation was derived for the case of a single path channel. This implies that the received signal is simply an attenuated and delayed version of the transmitted signal. However, in a multipath channel, the received signal is a superposition of several attenuated and delayed replicas of the transmit signal as shown in Figure 2.3. In such channels, neglecting non-linear effects, the received signal is modeled as:

$$r_B(t) = s_A(t) \star h(t) + n(t), \tag{2.11}$$

where $h(t)$ is the impulse response of the channel, and $\star$ denotes convolution. The impulse response is typically modeled as a series of impulses:

$$h(t) = \sum_{k=0}^{L-1} \alpha_k \delta(t - \tau_k),$$

where α_k and τ_k are the amplitude and TOA of the kth multipath component, $k = 0, 1, 2, \cdots, L$. For the transmit signal defined in (2.2), the received signal is given by

$$r_B(t) = s(t) \star h(t) + n(t) = \sum_{k=1}^{N_m} \sum_{l=0}^{L-1} \alpha_l p(t - kT_f - \tau_l) + n(t). \tag{2.12}$$

In this case, our goal is to estimate the TOA of the earliest arriving multipath component or *first-path*, as this corresponds to the propagation delay of the LOS path. From the above equation, and Figure 2.3, we see that in a multipath channel, the maximization of the correlation between the received signal and the transmit pulse may not yield an accurate estimate of the TOA of the first path. This is because multiple replicas of the transmitted pulse can overlap and shift the position of the correlation peaks. Therefore, the goal is to estimate τ_0 given $r_B(t)$ defined in (2.12), in the absence of the knowledge of $\{\alpha_l\}$, $l = 0, 1, \cdots, L-1$, and $\{\tau_l\}$, $l = 1, \cdots, L-1$. These unknown parameters are treated as *nuisance parameters* [34], and the accuracy of range estimation in this scenario is shown in [17] to be worse than the single-path AWGN case as stated in the theorem below:

Theorem 2 (Range Estimation Bound in AWGN with multipath). *The CRLB for the estimation of range based on TOA of a single path in a multipath environment in the presence of AWGN is greater than the CRLB of range estimation with an isolated single path in AWGN:*

$$\sigma^2_{r,\ multipath} > \sigma^2_{r,\ AWGN} \tag{2.13}$$

A closed-form expression for the CRLB in terms of the nuisance parameters [17] is hard to derive. A simplistic model for the effect of the multipath channel would be to assume that the transmitted pulse width increases by the delay spread of the channel without pulse distortion. Neglecting non-linear effects, we can then model the CRLB in the multipath case using

$$\sigma^2_{r,\ multipath} \approx \frac{K_R}{\xi} \tag{2.14}$$

where $K_R > K'_R$. The larger proportionality constant K_R in this case incorporates the observations that (i) we are trying to estimate the TOA of the earliest (LOS) path in the presence of other multipath, (ii) the performance of the TOA estimator depends on the SNR of the received signal. Consequently, for our purposes, the CRLB for TOA estimation in AWGN with multipath is modeled using

$$\sigma^2_{r,\ multipath} \geq \frac{K_R R^\beta}{K_P P_T} = K_E R^\beta$$

where $K_E = \frac{K_R}{K_P P_T} > K'_E$.

Example 2: TOA-based Ranging with UWB in multipath and AWGN With the parameters given in Example 1, assume that the multipath channel has a delay spread of approximately 5 ns (a typical delay spread value in LOS UWB multipath channels), and the received pulse is a broadened Gaussian pulse of width $T_w = 5$ ns (this is 10 times the value of T_w used in Example 1). In this case, the effective RMS bandwidth $B_e = 135$ MHz (smaller by a factor of 10 than the B_e in Example 1), and the corresponding value of K_E is 1.25×10^{-4}. At this value of K_E, the minimum standard deviation of an unbiased range estimate predicted by (2.10) at a distance $R = 10$ meters under LOS conditions is 11.2 centimeters.

2.2.3 Multiple Access Interference

In a multi-user environment, signals from simultaneously transmitting nodes interfere with each other and degrade the performance of TOA estimation. Assuming the presence of a large number of simultaneous users, using well-designed TH (or DS) codes, we can apply the Gaussian approximation [35], [37] for the multiple-access interference (MAI). Based on this assumption, we can further approximate the model for the CRLB for TOA-based range estimation in a multipath channel in the presence of MAI:

$$\sigma_r^2 \geq \frac{K_R}{\xi'} \tag{2.15}$$

where ξ' is the effective signal-to-interference-and-noise-ratio (SINR).

2.2.4 NLOS Propagation

In dense multipath propagation environments, especially indoors or in urban scenarios, the LOS path between nodes may be obstructed. When the direct path between two nodes is blocked, only reflections of the UWB pulse from scatterers reach the receiving node as shown in Figure 2.4. In this case, the delay of the earliest arriving path does not represent the true TOA of a path corresponding to the LOS distance. Since the signal travels an distance in excess of the LOS distance, a positive bias called the *NLOS bias error* is introduced in the TOA-based range estimate.

NLOS propagation has popularly been classified as "soft" NLOS, where the LOS multipath component is detectable albeit attenuated, and "hard" NLOS, where the LOS path is severely attenuated or absent. From a localization standpoint, the soft-NLOS cases are classified as LOS scenarios, since the TOA of the LOS multipath component can still be estimated for ranging purposes, i.e., range estimates are not necessarily biased in soft-NLOS propagation environments. The NLOS problem is investigated in detail in Chapter 4.

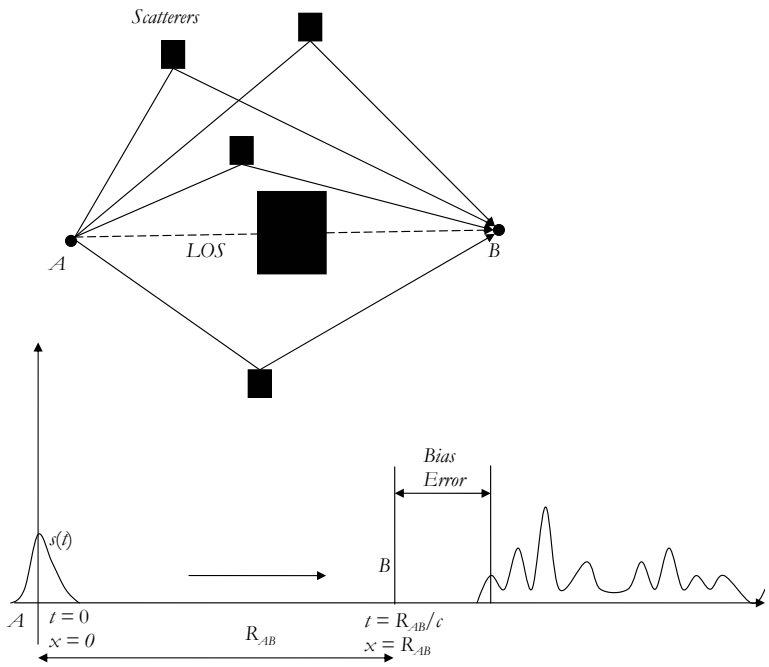


Figure 2.4: Illustration of the effect of NLOS Propagation on TOA-based range estimation

2.3 A two-way packet handshake for distributed ranging

The potential range measurement accuracy with UWB signals, *given perfect synchronization between nodes* is very high, as demonstrated in the previous section. This implies that the synchronization between the nodes becomes an important factor affecting TOA estimation accuracy. In the absence of a common clock between the nodes, the round-trip time (RTT) of a signal transmitted between two nodes can be measured to estimate the distance between two nodes.

A mechanism for measuring the RTT, and thereby the distance between asynchronous nodes of a network, comprises a packet-handshake. Suppose node A of the network would like to know the range R_{AB} to node B, as shown in Figure 2.5. Node A first transmits a packet, termed the "range-initiate" (RI) packet, to node B at time $t = t_1$. This packet reaches B at $t = t_1 + \tau_0$. Node B synchronizes itself to the TOA of the received packet. Node B then sends a response packet, called a "range-response" (RR) packet, after a certain known delay T. Node A receives the response packet and determines the TOA of the packet which is $t_2 = 2\tau_0 + t_1 + T$. Node A then computes the difference $\Delta t = t_2 - t_1 = 2\tau_0 + T$, from which the known delay T can be subtracted. The range R_{AB} can be computed using $R_{AB} = \frac{c(t_2 - t_1 - T)}{2}$, where c is the speed of light. A useful feature of this scheme is that any

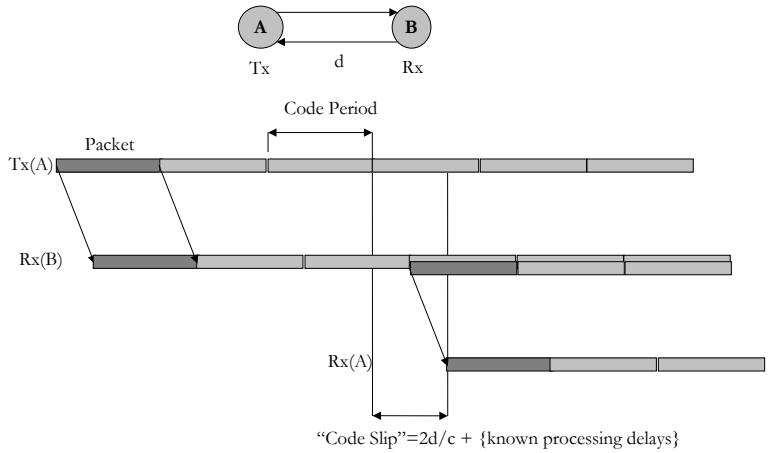


Figure 2.5: Two-Way Ranging Mechanism

data can be exchanged in the data portion of the packet headers, as only the TOAs of the packets are used for ranging purposes.

The two-way mechanism described above essentially allows the two nodes to synchronize their clocks, thereby creating a common time-reference for ranging. The above equations assume the absence of processing delays, but these can be eliminated via calibration. Two important issues that impact the performance of the scheme described above are: (i) the accuracy of the packet TOA estimation (also called leading-edge-detection (LED)) in the presence of noise, and (b) the clock drift between the nodes.

The problem of leading-edge or first-path detection for UWB signals is related to the problem of fine acquisition. As mentioned previously, in order to compute the range accurately, we need to determine the TOA τ_0 of the first path of the received multipath profile, as this corresponds to the propagation delay of the signal between the two nodes. Coarse acquisition algorithms typically "lock" onto the strongest multipath components in the received profile. For LOS multipath profiles, the acquisition circuit typically locks onto the LOS path, and the leading edge may be estimated using the received pulse-shape as shown in Figure 2.6. However, this procedure is not straightforward for the NLOS [1] case since the LOS path is either absent or severely attenuated. An error of 3 nanoseconds in the TOA estimate results in a range estimate error of approximately 1 meter. Therefore, specific algorithms need to be designed to search for the first-path in order to estimate τ_0 accurately, especially at low SNR. These algorithms are then required to be applied at both stages of the two-way packet exchange described above in order to guarantee high range measurement accuracy.

The other important issue that affects the performance of the described packet handshake

[1] The figure refers to the soft-NLOS case where the LOS path is present albeit attenuated. The hard-NLOS case, where the LOS path is absent, is the subject of Chapter 4.

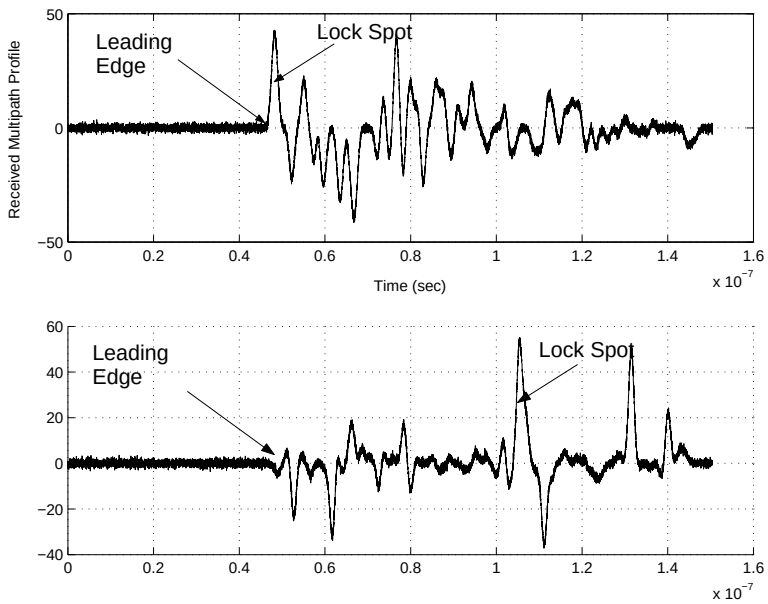


Figure 2.6: Typical leading edges and acquisition lock-spots for LOS and NLOS multipath profiles

mechanism is clock drift. Suppose nodes A and B have uniform normalized clock drift rates δ_A and δ_B respectively. This implies that the delay T' after receiving the RI packet from A at which B transmits an RR packet is given by $T' = T(1+\delta_B)$. Then node A receives the RR packet at $t_2 = t_1 + 2\tau_0 + T'$, which it perceives as $t_2' = t_1 + (2\tau_0 + T')(1+\delta_A)$. Therefore,

$$R_{AB} = \frac{c(t_2' - t_1 - T)}{2} = c\tau_0\left(1+\delta_A\right) + \frac{cT}{2}\left(\delta_A + \delta_B + \delta_A\delta_B\right).$$

This implies that the range error is given by:

$$\epsilon_R = c\tau_0\delta_A + \frac{cT}{2}\left(\delta_A + \delta_B + \delta_A\delta_B\right). \tag{2.16}$$

As an example, let $\delta_A = \delta_B = +1$ parts-per-billion (ppb). Further suppose $T = 1$ ms and $\tau_0 = 10$ ns. With these values, the range estimation error due to clock drift is given by 0.3 meters. When $\delta_A = +1$ ppb, and $\delta_B = -1$ ppb then the effects of the clock drift on the range error tend to cancel each other out. From the above example, we see that the range error is very sensitive to the variation of the clock drifts of the nodes. In this work, we ignore the effect of clock jitter and drift on range-estimation. The effect of the combination of clock drift, jitter and noise on the ranging error has been discussed in [38].

In the following section, we discuss a simple non-linear leading-edge detection scheme based on thresholding the energy of samples from the received multipath profile. It must be emphasized that this development was done prior to, and independently of [39], [40], [41].

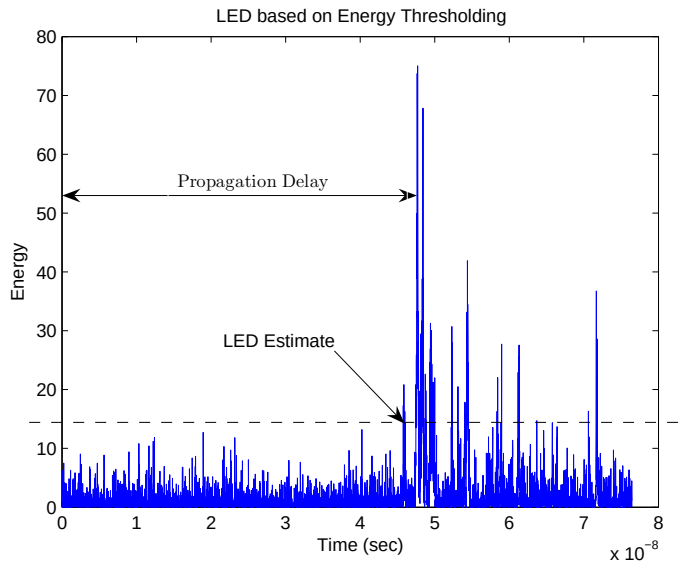


Figure 2.7: Thresholding the energy of the received signal samples: the threshold α is set based on the energy of noise samples.

2.4 A nonlinear energy - thresholding estimator for TOA - based range estimation

Given the received multipath profile, a direct approach to finding the propagation delay τ_0 is to search for the "first indication of signal energy". Suppose the observed multipath profile at the receiver after averaging of several received pulse responses (in order to improve the SNR) is given by:

$$r(t) = \sum_{l=0}^{L} \alpha_l p(t - \tau_l) + n(t).$$

The sampled version of the above signal is given by

$$x[k] = \underbrace{\sum_{l=0}^{L} \alpha_l p(kT_s - \tau_l)}_{y[k]} + z[k], \tag{2.17}$$

where T_s is the sampling interval, $x[k] = r(kT_s)$, and $z[k] = z(kT_s) \sim \mathcal{N}(0, \sigma_N^2)$ are uncorrelated Gaussian noise samples.

We can then obtain the multipath energy profile by squaring $x[k]$ as shown in Figure 2.7. From Figure 2.7, we see that there are two distinct regions: a "noise region" and a "signal region". The key idea is to set a threshold for the energy which is crossed only at the

transition between the noise and signal regions. If the threshold is crossed in the noise region, this constitutes a *early-detection*, and results in $\hat{\tau}_0 < \tau_0$. If the threshold is crossed in the signal region after the arrival of the first path, this constitutes a *late-detection* and results in $\hat{\tau}_0 > \tau_0$. We also assume that we can estimate the noise variance σ_N^2 from samples in the noise region. The threshold is selected to be a factor α times the noise variance estimated from the noise region.

Assuming $N_p = \lfloor \frac{\tau_0}{T_s} \rfloor$ represents the time-index corresponding to the TOA of the first-path in $y[k]$, the instantaneous energy values are given by:

$$Z_k = x^2[k] = \begin{cases} z^2[k], & 0 \leq k \leq N_p, \\ (y[k] + z[k])^2, & k \geq N_p. \end{cases}$$

Since the noise samples $z[k]$ are uncorrelated zero-mean Gaussian random variables, the Z_k can be modeled as a χ^2 random variable with one degree of freedom (χ_1^2) if it is a noise sample and if it contains signal energy, as a non-central χ^2 random variable with one degree of freedom $\bar{\chi}_1^2(y[k])$, where $y[k]$ is the non-centrality parameter. Explicitly, for $k < N_p$, $Z_k \sim \chi_1^2$, and for $k \geq N_p$, $Z_k \sim \bar{\chi}_1^2(y[k])$. Therefore, the probability density function of Z_k, in the noise region $0 \leq k \leq N_p$ is defined by

$$f_{Z_k}(z) = \frac{e^{-\frac{z}{2\sigma_N^2}}}{\sqrt{2\pi\sigma_N^2 z}}, \quad 0 \leq k \leq N_p.$$

Assuming a threshold $Z_0 = \alpha\sigma_N^2$, the probability that Z_k is greater than Z_0 for $0 \leq k \leq N_p$ is given by

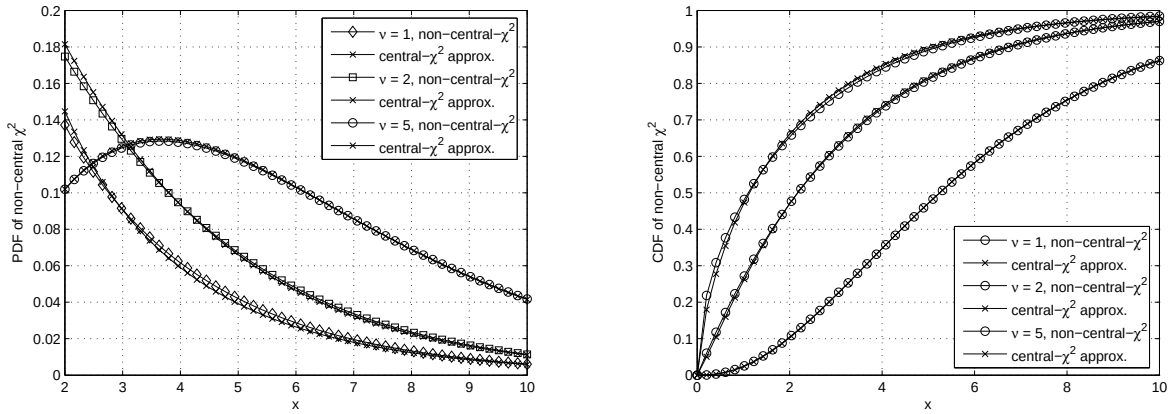
$$\begin{aligned} p_k = \Pr\left[Z_k \geq Z_0\right] &= 1 - F_{Z_k}(Z_0) = 1 - \int_0^{\frac{Z_0}{2\sigma_N^2}} \frac{e^{-\frac{x}{2}}}{\sqrt{2\pi x}} dx \\ &= 1 - \gamma_{inc}\left(\frac{\alpha}{2}, \frac{1}{2}\right), \quad 0 \leq k \leq N_p, \end{aligned} \tag{2.18}$$

where $\gamma_{inc}(v, a)$ is the incomplete gamma function defined as:

$$\gamma_{inc}(v, a) \triangleq \frac{1}{\Gamma(a)} \int_0^v t^{a-1} e^{-t} dt.$$

For a given α, this probability is identical for all $0 \leq k \leq N_p$. For $k \geq N_p$, Z_k is a non-central χ^2 random variable with one degree of freedom with centrality parameter $y[k]$. Using the result from [42], the non-central χ^2 distribution with ν degrees of freedom and a non-centrality parameter λ_k is approximated by the central χ^2 distribution with ν^* degrees of freedom:

$$f_{\bar{\chi}^2}(x|\ \nu, \lambda_k) \approx \frac{1}{c} f_{\chi^2}\left(\frac{x}{c}\middle|\ \nu^*\right), \quad c = \frac{\nu + 2\lambda_k}{\nu + \lambda_k},\ \nu^* = \frac{(\nu + \lambda_k)^2}{\nu + 2\lambda_k}. \tag{2.19}$$



(a) (b)

Figure 2.8: On the validity of the non-central χ^2 approximation: (a) Comparison of the non-central density functions with the approximate density functions for $\nu = 1,\ 2$ and 5. (b) Comparison of the non-central CDFs with the approximate CDFs for $\nu = 1,\ 2$ and 5.

The validity of this approximation in terms of the probability density function and the cumulative distribution function is verified in Figures 2.8(a) and 2.8(b) for different degrees of freedom. Using this approximation,

$$\begin{aligned} p_k &= \Pr\left[Z_k \geq Z_0\right] = 1 - \gamma_{inc}\left(\frac{\alpha}{2c_k}, \frac{\nu_k^*}{2}\right), \\ c_k &= \frac{1+2\lambda_k}{1+\lambda_k},\ \nu_k^* = \frac{(1+\lambda_k)^2}{(1+2\lambda_k)},\ \lambda_k = y^2[k],\ k \geq N_p. \end{aligned} \tag{2.20}$$

Equations (2.18) and (2.20) define p_k for all values of k. Figure 2.9 shows p_k versus k for a Gaussian LOS pulse at an SNR of 15 dB with threshold $\alpha = 10$.

The probability that $\hat{\tau}_0 = kT_s$ is the probability that Z_k crosses the threshold Z_0, given that none of the $(k-1)$ previous samples crossed the threshold. Therefore, if P_k denotes the probability that the kth sample is picked as the index corresponding to the TOA of the first-path,

$$P_k = \left(\prod_{j=1}^{k-1}(1-p_j)\right) p_k.$$

The probability of picking none of the samples (a *missed detection*, resulting in the default $\hat{\tau}_0 = 0$) is given by

$$P_0 = \prod_j (1-p_j). \tag{2.21}$$

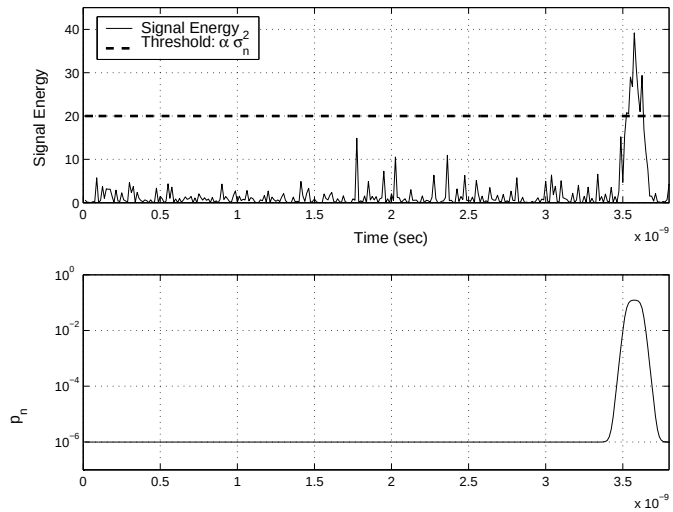


Figure 2.9: The probability p_k for a given multipath profile: (top) A realization of the noisy received signal $x[k]$, (bottom) The probability p_k of energy values $x^2[k]$ crossing the threshold $Z_0 = 10\sigma_N^2$.

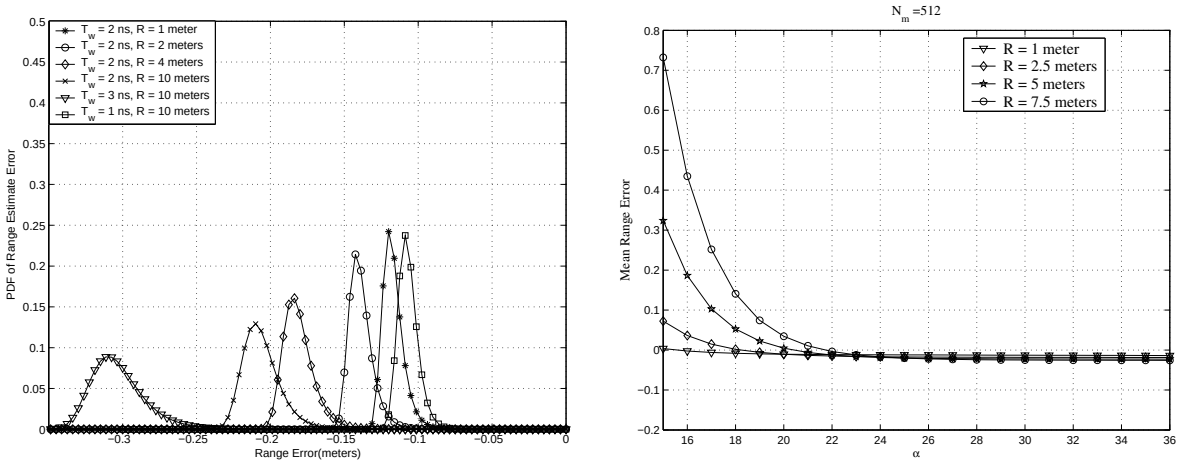
Then, the probability density function of the range estimate r is then given by

$$f_r(\rho) = \sum_k P_k \delta(\rho - ckT_s) = \sum_k p_k \left(\prod_{j=1}^{k-1} (1 - p_j) \right) \delta(\rho - ckT_s).$$

Denoting the ranging error by $\epsilon_R = r - R$, the probability density function of ϵ_R is given by

$$f_{\epsilon_R}(\rho) = \sum_k P_k \delta(\rho - ckT_s + R) = \sum_k p_k \left(\prod_{j=1}^{k-1} (1 - p_j) \right) \delta(\rho - ckT_s + R). \tag{2.22}$$

It is evident that the accuracy of range estimation depends on the sampling resolution of the received multipath profile. The envelope of the distribution $f_{\epsilon_R}(x)$ of the negative range error $-\epsilon_R$ for a single path AWGN channel is shown in Figure 2.10(a) for different distances R and pulse-widths T_w. We see that the density of the range error is (i) not symmetric, and (ii) has a negative bias. This is because (i) the probability of a time index being selected as the TOA is conditioned on none of the previous points being picked, leading to asymmetry, and (ii) the threshold value $Z_0 = \alpha\sigma_N^2$ is set so as to reduce the probability of false-alarm (which results in $r < R$) in the noise region, resulting in a negative bias. We note that the latter result is more pronounced at larger distances since in this case, the noise variance is larger. This is seen in Figure 2.10(b) where the variation of the mean negative range error with distance is shown. At larger distances, the reduction in SNR results in larger values of Z_0 relative to signal energies. This results in a higher probability of late detection

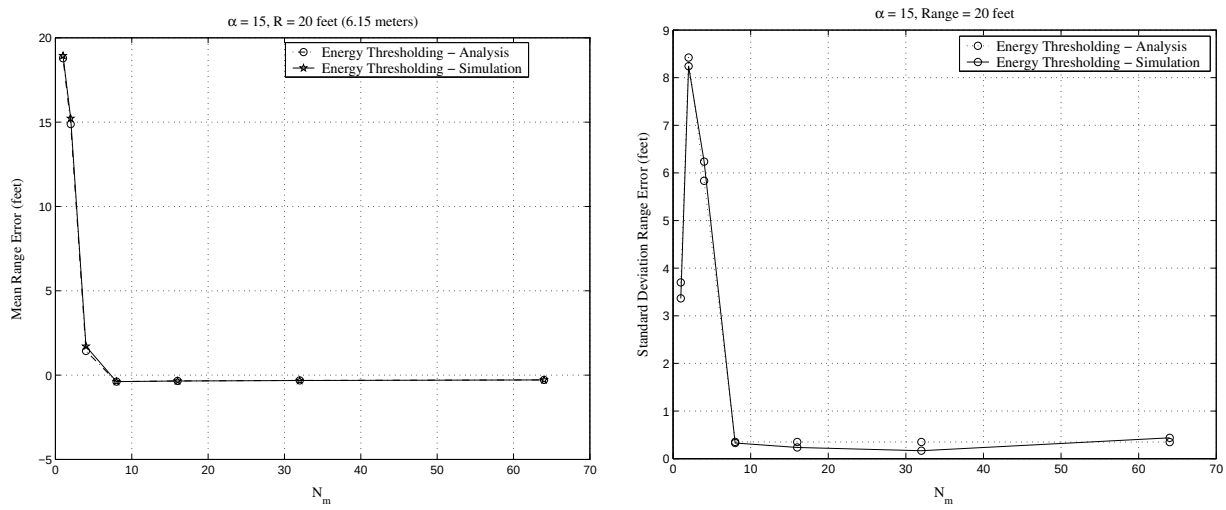


(a) (b)

Figure 2.10: The envelope of the distribution of the range error $R-r$ (or $-\epsilon_R$): (a) PDF of the negative range estimation error $R-r$ for different distances and the pulse widths. Increasing the true range (i.e. decreasing the SNR) and increasing the pulse-width T_w increase the mean and variance of $R-r$. (b) Variation of the mean of the negative range error with the threshold α for different ranges.

as the values R or α increase, leading to a positively biased ϵ_R. This effect is observed in Figure 2.10(b). The bias of the range estimates can be reduced using the calibration of range estimates.

Figures 2.11(a) and 2.11(b) show the comparison of the performance of the described energy thresholding scheme in terms of the mean and standard deviation of the range error obtained through analysis and simulation. We see that the analytical modeling of the range error in (2.22) closely matches the performance obtained by simulation. Figure 2.12 compares the performance of the thresholding method described for a single path AWGN case for different values of N_m and β. We see that for larger values of β, a larger value of N_m is required to compensate for the decrease in received SNR and the corresponding degradation in range estimation accuracy. For small values of R (high SNR), as $R \to 0$, we see that the accuracy of the described scheme saturates, approaching the CRLB. Further, it can be verified that the variance of the range estimates is approximately proportional to R^β for larger values of R (slope of the logarithm of the variance on the y-axis in Figure 2.12 is proportional to β), as predicted by (2.10).



(a) (b)

Figure 2.11: The comparison between analysis and simulation of energy-based TOA estimation: in terms of the (a) mean range error $(R - r)$, and (b) the standard deviation of the range error $(R - r)$. The threshold used was $\alpha = 15$, which implies a probability of false alarm of 5×10^{-2} if 100 noise samples are present.

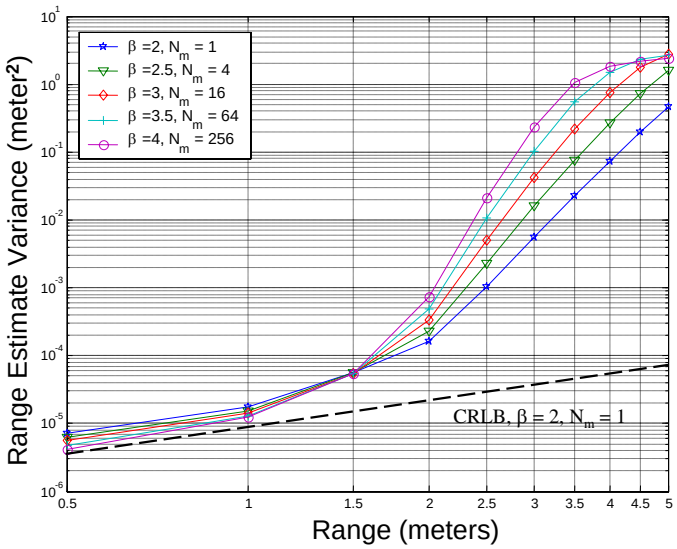
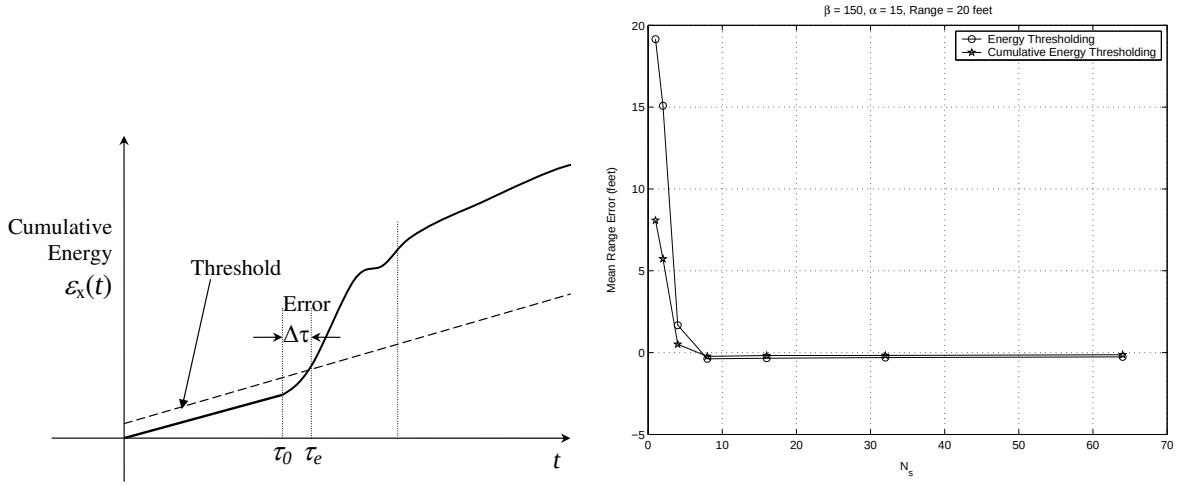


Figure 2.12: Variance of the range estimates obtained through energy thresholding versus the distance R for different values of β and N_m. The CRLB for $\beta = 2$, $N_m = 1$ is also shown. The legend is repeated here for clarity: (1) Blue Star: $\beta = 2$, $N_m = 1$, (2) Green Triangle: $\beta = 2.5$, $N_m = 4$, (3) Red Diamond: $\beta = 3$, $N_m = 16$, (4) Cyan Plus: $\beta = 3.5$, $N_m = 64$, (5) Magenta Circle, $\beta = 4$, $N_m = 256$. The x-axis runs from 0.5 meters to 5 meters, and the y-axis runs from 10^{-6} to 10 meter^2.



(a) (b)

Figure 2.13: (a) Illustration of cumulative energy thresholding, (b) Comparison of the mean range error $(R - r)$ obtained using the energy thresholding and cumulative energy thresholding methods.

2.4.1 Cumulative Energy Thresholding

In this case, the values of interest are the cumulative received signal energy values:

$$Z_k = \sum_{j=0}^{k} x^2[j] = \begin{cases} \sum_{j=0}^{k} z^2[j], & 0 \leq k \leq N_p, \\ \sum_{j=0}^{k} (y[j] + z[j])^2, & k \geq N_p. \end{cases}$$

Once again, the variance of the squared noise samples σ_N^2 is estimated from the noise portion of the received multipath profile. The threshold in this case is defined as:

$$Z_{T,k} = \sigma_N^2 (k + b), \tag{2.23}$$

where b is an offset parameter. The foundation for such a threshold is that as the noise energy samples are uncorrelated χ^2 random variables with mean σ_N^2, the average cumulative energy of noise samples increases linearly as $\sigma_N^2 k$. However, the cumulative energy increases at a much higher rate when the signal energy is included as seen in Figure 2.13(a). With an appropriate choose of b, this transition in cumulative signal energy can be captured by the linearly increasing threshold defined in (2.23). Figure 2.13 compares the performance of the energy thresholding and cumulative energy thresholding techniques in terms of the mean range error $R - r$. We see that the cumulative energy thresholding method generates more accurate range estimates in terms of mean range error than the energy thresholding

Table 2.1: LOS and NLOS Range Measurement Error Statistics

	LOS	**NLOS**
Number of Measurements	260	623
Mean Error $\langle r_i - R_i \rangle$	0.02 meters	1.90 meters
Standard Deviation of Error	0.23 meters	2.09 meters
$\left\langle \left(\frac{r_i}{R_i} - 1 \right)^2 \right\rangle$	0.005	0.15
Kurtosis	4.74	15.66

technique. These gains result from the added averaging arising from the cumulative sum of energies that reduces the probability of early detection.

2.5 Range Error Measurements and Modeling

Thus far, we have discussed bounds on the accuracy of range estimation, and the performance of practical estimators in single path AWGN channels. In this section, we discuss the performance of practical range estimation in measured multipath channels. Table 2.1 summarizes some of the parameters observed in the ranges estimated from indoor measurements [1] over distances of 1-30 meters. The range estimates are obtained using the energy thresholding TOA estimator applied to indoor received signal measurements taken with Bicone antennas [1] over distances 1-30 meters. The measurement setup and other details can be found in [1] and in Chapter 6. The range estimates were calibrated using a reference measurement at 1 meter, and an energy threshold $\alpha = 36$ was used.

In the case of unbiased range estimates, an estimate of K_E can be estimated using

$$\hat{K_E} = \frac{\epsilon_i^2}{R_i^2} = \frac{(r_i - R_i)^2}{R_i^2} = \left(\frac{r_i}{R_i} - 1 \right)^2 . \tag{2.24}$$

We see from Table 2.1 and Figure 2.14 that a rough estimate for the value of K_E in LOS scenarios is about 0.005. At a distance $R = 10$ meters, this implies a standard deviation of range error on the order of 0.5 meters.

Figure 2.15 shows the empirical histograms of the range error $(r_i - R_i)$ from measurements. We see that LOS range estimates after calibration are approximately unbiased, symmetric, and have a small variance. On the other hand, the NLOS range estimates have a large positive bias with high probability and this bias is much larger than the variance of the range estimates in the LOS case (this is exploited in the mitigation of the unknown bias errors in Chapter 4). In view of the need for a mathematically tractable model for the

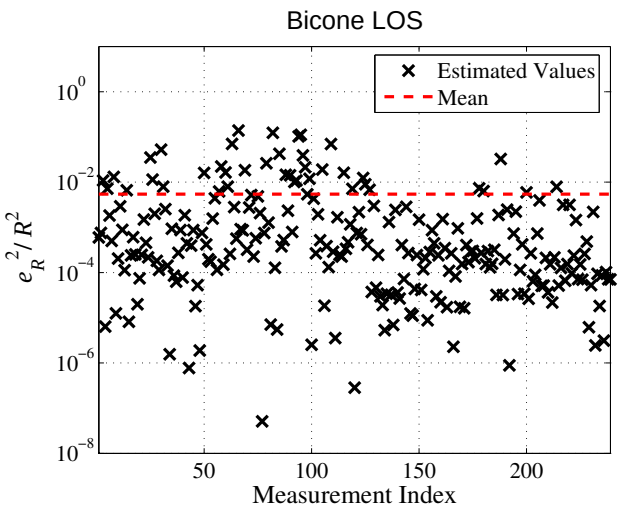
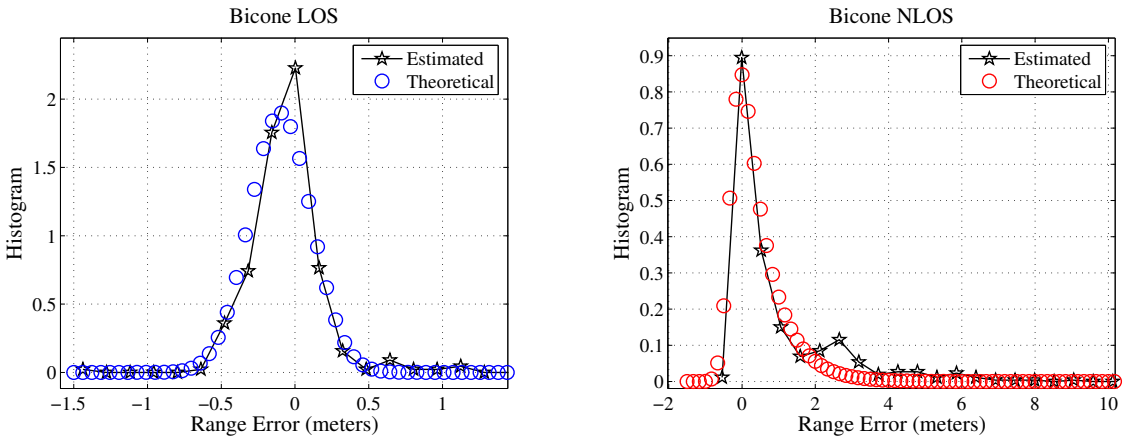


Figure 2.14: Estimation of K_E



(a) (b)

Figure 2.15: Empirical histograms of the measured range estimates in (a) LOS, and (b) NLOS measurement scenarios.

accuracy of TOA-based range estimation, we use the following model for range information, based on the above observations:

$$r_i \sim \begin{cases} \mathcal{N}(R_i, \sigma_i^2), & \sigma_i^2 = K_E R_i^{\beta_L}, \quad \text{LOS Range Estimates,} \\ \mathcal{N}(b_i + R_i, \sigma_i^2), & \sigma_i^2 = K_E R_i^{\beta_N}, \quad \text{NLOS Range Estimates,} \end{cases} \tag{2.25}$$

where β_L and β_N are the path-loss exponents in LOS and NLOS cases respectively. The assumption that LOS range estimates are Gaussian distributed is fairly common in the literature on position-location [9], [19], although experimental results [43] indicate that range estimates may not be Gaussian in the presence of dense multipath. It is further assumed that in the LOS case, errors due to hardware, clock jitter, processing delays, etc., are eliminated via calibration.

As discussed previously, the NLOS range estimates are positively biased with high probability. When the distribution of the bias errors b_i is unknown *a priori*, they can be modeled as uniformly distributed random variables: $b_i \sim \mathcal{U}(0, B_{\max})$, where $B_{\max}$ is the maximum bias. Measurement results [44] suggest that the NLOS bias errors may be exponentially distributed in dense multipath channels, which has also been observed in indoor UWB NLOS measurements [1], [43]. The model for r_i in NLOS conditions is therefore the superposition of an unbiased Gaussian measurement noise and a random positive bias:

$$r_i = R_i + n_i + b_i,$$

where b_i is exponentially distributed with parameter λ_b, and is assumed to be independent of the range measurement noise. The rationale for the bias being exponentially distributed is that the arrival of multipath components can be modeled using a Poisson process [13], which implies exponentially distributed inter-arrival times [1]. Therefore, in NLOS scenarios, if the LOS path is absent or attenuated, and a path other than the LOS path is selected for the estimation of $\hat{\tau}_0$, this is likely to result in an exponentially distributed bias error. Based on these assumptions, as shown in Appendix 2A, the probability density function of r_i can be written as:

$$f_{r_i}(\rho\,|\text{NLOS}) = \frac{1}{2\lambda_b} \exp\left(\frac{\sigma_i^2}{2\lambda_b^2}\right) \exp\left(-\frac{(\rho - R_i)}{\lambda_b}\right) \left[1 + \operatorname{erf}\left(\frac{\lambda_b(\rho - R_i) - \sigma_i^2}{\sqrt{2}\sigma_i\lambda_b}\right)\right].$$

To summarize the statistical models for range estimates in LOS and NLOS scenarios:

$$f_{r_i}(\rho) = \begin{cases} (2\pi\sigma_i^2)^{-1/2} \exp\left(-\frac{(\rho - R_i)^2}{2\sigma_i^2}\right), & \text{LOS,} \\ \frac{1}{2\lambda_b} \exp\left(\frac{\sigma_i^2}{2\lambda_b^2}\right) \exp\left(-\frac{(\rho - R_i)}{\lambda_b}\right) \left[1 + \operatorname{erf}\left(\frac{\lambda_b(\rho - R_i) - \sigma_i^2}{\sqrt{2}\sigma_i\lambda_b}\right)\right], & \text{NLOS.} \end{cases} \tag{2.26}$$

Figure 2.15 compares the histogram of the range error $(r_i - R_i)$ from measurements, with the corresponding statistical range models in (2.26) for the LOS and NLOS cases. We see that although the number of range measurements is small, the statistical models for the range errors based on the parameters in Table 2.1 are reasonably close to the observed

histograms. In the rest of this dissertation, the model for range estimation in LOS and NLOS scenarios given by (2.25) is utilized. Based on this model for range estimates, in the following chapter, we characterize bounds on the localization accuracy of nodes given sufficient range information.

2.6 Conclusions

In this chapter, we provided a characterization of the performance of the efficient estimator for synchronized TOA-based ranging in AWGN and in the presence of multipath, through the derivation of the Cramer-Rao Lower Bound for the variance of an unbiased range estimator. We then described a two-way packet handshake for distributed ranging in the absence of synchronization between nodes. A practical non-efficient estimator based on energy thresholding of the received multipath profile was then analyzed in detail for single path AWGN channels. We then characterized the performance of practical range estimators in measured indoor LOS and NLOS multipath channels, in order to derive a model for LOS and NLOS range estimates that is used in the rest of this dissertation.

2.7 Appendix 2A: Distribution of r_i under NLOS conditions

Under NLOS propagation conditions, the range estimate is modeled as a superposition of unbiased Gaussian range measurement noise, and an exponentially distributed bias error:

$$r_i = R_i + n_i + b_i,$$

where n_i is a zero-mean Gaussian random variable with variance σ_i^2, and b_i is exponentially distributed with mean λ_b. Therefore, the probability density function of r_i is defined by a convolution of the density functions of n_i and b_i:

$$\begin{aligned}
f_{r_i}(\rho\,|\mathrm{NLOS}) &= \left(2\pi\lambda_b^2\sigma_i^2\right)^{-1/2}\int_0^{\infty}\exp\left(-\frac{(\rho-R_i-b)^2}{2\sigma_i^2}\right)\exp\left(-\frac{b}{\lambda_b}\right)db,\\
&= \left(\pi\lambda_b^2\right)^{-1/2}\int_{-\infty}^{\frac{r-R_i}{\sqrt{2}\sigma_i}}\exp\left(-y^2\right)\exp\left(-\frac{\rho-R_i-\sqrt{2}\sigma_i y}{\lambda_b}\right)dy,\\
&= \left(\pi\lambda_b^2\right)^{-1/2}\exp\left(\frac{\sigma_i^2}{2\lambda_b^2}\right)\exp\left(-\frac{\rho-R_i}{\lambda_b}\right)\int_{-\infty}^{\frac{\rho-R_i}{\sqrt{2}\sigma_i}}\exp\left(-\left(y-\frac{\sigma_i}{\sqrt{2}\lambda_b}\right)^2\right)dy,\\
&= \frac{1}{2\lambda_b}\exp\left(\frac{\sigma_i^2}{2\lambda_b^2}\right)\exp\left(-\frac{(\rho-R_i)}{\lambda_b}\right)\left[1+\mathrm{erf}\left(\frac{\lambda_b(\rho-R_i)-\sigma_i^2}{\sqrt{2}\sigma_i\lambda_b}\right)\right],
\end{aligned}$$

where $\mathrm{erf}\,(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$. The mean and standard deviation of the range estimate are respectively given by:

$$\begin{aligned} \mu_{r_i} &= R_i + \lambda_b, \\ \sigma_{r_i} &= \sqrt{\sigma_i^2 + \lambda_b^2}. \end{aligned}$$

Chapter 3

Bounds on Location Estimation

3.1 Introduction

In the previous chapter, we derived a model for the range information available to an unlocalized node that can be used to estimate its location. In ad hoc UWB PoLoNets, these range estimates may be available to an unlocalized node from either fixed anchors or localized reference nodes, henceforth referred to simply as "anchors". Given range estimates from anchors, and the corresponding anchor locations, the location of an unlocalized node can be estimated.

In this chapter, a systematic formal study of the properties of the CRLB on the location estimation given unbiased Gaussian range estimates from anchors is presented. The dependence of localization accuracy on the accuracy of range estimates, the impact of node geometry and the number of range estimates and other parameters is analytically characterized. These insights into the properties of the CRLB are used in establishing the connection between localization accuracy and multiple access schemes and power-control algorithms in Chapters 7 and 9. We then compare the performance of various practical estimators based on the Least-squares (LS) formulation with the CRLB in order to (i) verify that similar trends were observed in terms of the properties, and (ii) quantify the difference between the CRLB and practical estimators.

Previous Work included in this Chapter: CRLB for location estimation with unbiased Gaussian range estimates [9], [19], the Least-Squares estimator [20], Iterative refinement using Levenberg-Marquhardt iteration [21].

Original Contributions in this Chapter:

- The systematic study of the properties of the CRLB provides insights into the factors than impact localization accuracy. A novel characterization of the geometric dilution of precision was developed, which allowed us to individually assess the impact of the variance of range estimates, geometry of anchors, and the number of range estimates on average localization accuracy.
- Similar trends were observed with practical estimators and are used in establishing the connection between localization accuracy and multiple access schemes and power-control algorithms [22], [23].

The organization of this chapter is as follows: Section 3.2 discusses the CRLB on location estimation with unbiased Gaussian range estimates from anchors. In Section 3.3, we present a systematic analysis of the properties of the CRLB in terms of the accuracy of range estimates, and the geometry and number of anchors. Section 3.4 describes the formulation of practical location estimation based on a Least-Squares (LS) approach. The modification of the LS estimator to incorporate weighting of range estimates is discussed in Section 3.5. Section 3.6 discusses the use of iterative refinement of LS location estimates in order to achieve gains in average localization accuracy. This chapter concludes in Section 3.7.

3.2 Cramer-Rao Lower Bound on Location Estimation

In this section, we derive bounds on the performance of location estimators given unbiased Gaussian range estimates that are defined in (2.25). Let us assume that we are trying to estimate the location $\mathbf{x} = [x\ y]^T$ of a node given m unbiased Gaussian range estimates $r_i \sim \mathcal{N}(R_i, \sigma_i^2)$, from anchors with known locations $\mathbf{x}_i$, $i = 1, 2, \cdots, m$. Since the available range estimates are noisy, the exact solution $\mathbf{x}$ cannot be determined in general. The error in the location estimate is denoted by $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$, where $\hat{\mathbf{x}} = [\hat{x}\ \hat{y}]^T$ represents the estimate of the true node location $\mathbf{x}$. The covariance matrix of the error is given by:

$$\mathbf{C}_{\hat{\mathbf{x}}} = E\left\{\mathbf{e}\mathbf{e}^T\right\} = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix}.$$

The CRLB for the estimation of a node's location given unbiased Gaussian range estimates from known locations has been derived previously in [9], [19], [45]. In terms of the error covariance matrix,

$$\mathbf{C}_{\hat{\mathbf{x}}} \geq \mathbf{I}^{-1} \Rightarrow \mathbf{C}_{\hat{\mathbf{x}}} - \mathbf{I}^{-1} \geq \mathbf{0}, \tag{3.1}$$

where $\mathbf{I}$ is the Fisher information matrix:

$$\mathbf{I} = \begin{bmatrix} \sum_{i=1}^{m} \frac{\cos^2(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} \\ \sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^{m} \frac{\sin^2(\alpha_i)}{\sigma_i^2} \end{bmatrix}, \tag{3.2}$$

where α_i is the orientation (angle) of the ith anchor node relative to the node whose location is being estimated as shown in Figure 3.1. The proof of the above equation is given in Appendix 3A. The term that we use to quantify the accuracy of node's location estimate is called the *localization error* and is defined by

$$\mathbf{e}^T\mathbf{e} = \|\mathbf{x} - \hat{\mathbf{x}}\|^2. \tag{3.3}$$

The *root localization error* is defined as the Euclidean distance between $\mathbf{x}$ and $\hat{\mathbf{x}}$: $\sqrt{\mathbf{e}^T\mathbf{e}} = \|\mathbf{x} - \hat{\mathbf{x}}\|$. The *average localization error* is defined as:

$$\Omega_{\mathbf{x}} = E\left\{\|\mathbf{x} - \hat{\mathbf{x}}\|^2\right\} = E\left\{\mathbf{e}^T\mathbf{e}\right\}. \tag{3.4}$$

Assuming the x and y coordinates of the node and its estimate are independent, and since the CRLB bounds the performance of an *unbiased estimator*, the average localization error for CRLB analysis can be written in terms of the variance of the location estimate $\mathbf{x}$ in the x and y coordinates:

$$\Omega_{\mathbf{x}} = \mathrm{Tr}\left\{\mathbf{C}_{\hat{\mathbf{x}}}\right\} = E\left\{(x - \hat{x})^2\right\} + E\left\{(y - \hat{y})^2\right\} = \sigma_x^2 + \sigma_y^2.$$

The average localization error defined above is a convenient scalar means of quantifying the CRLB defined in (3.1). The CRLB in terms of the average localization error $\Omega_{\mathbf{x}}$ is given by [19]:

$$E\left\{\mathbf{e}^T\mathbf{e}\right\} \geq \Omega_{\mathbf{x},CRLB} = \frac{\sum_{i=1}^{m} \frac{1}{\sigma_i^2}}{\sum_{i=1}^{m} \sum_{j=1,j>i}^{m} \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2 \sigma_j^2}}. \tag{3.5}$$

The derivation of the above equation is given in Appendix 3A. From (3.5), the localization error is, in general, a function of (i) the number of range estimates (m), (ii) the accuracy of the range estimates (σ_i^2, $i = 1, 2, \cdots, m$) and (iii) the geometry of anchor nodes (α_i, $i = 1, 2, \cdots, m$). For the UWB signals modeled in this work, from (2.25) and (3.5),

$$\Omega_{\mathbf{x},CRLB} = \frac{K_R \sum_{i=1}^{m} \frac{1}{R_i^{\beta}}}{K_P P_T \sum_{i=1}^{m} \sum_{j=1,j>i}^{m} \frac{\sin^2(\alpha_i - \alpha_j)}{R_i^{\beta} R_j^{\beta}}}. \tag{3.6}$$

For simplicity, we adhere to the notation used in (3.5) and present a detailed analysis of the properties of the CRLB in the following section.

3.3 Properties of the CRLB

In this section, we look at the impact of the parameters detailed above on the average localization error $\Omega_{\mathbf{x},CRLB}$ corresponding to the CRLB. In the following, the results are presented as a series of theorems and corollaries, with the salient observations summarized at the end of the section.

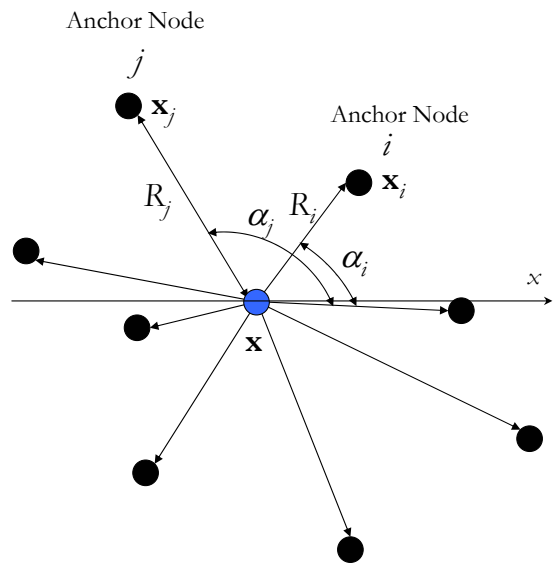


Figure 3.1: Effect of the Geometry of nodes on the localization error.

In order to isolate the impact of geometry from the impact of range estimate variances on the average localization error, we define the *generalized Geometric Dilution of Precision* (GGDOP) as follows:

Definition 1 (Generalized Geometric Dilution of Precision). *Suppose we define*

$$\gamma_m \triangleq \sum_{i=1}^{m} \frac{1}{\sigma_i^2} > 0, \;\; \psi_m \triangleq \sum_{i=1}^{m} \sum_{j=1, j>i}^{m} \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2 \sigma_j^2} \geq 0,$$

then the generalized Geometric Dilution of Precision (GGDOP) is defined by

$$\mathcal{G}_m = \frac{\gamma_m^2}{\psi_m} = \Gamma_m^{-1} \geq 0.$$

Γ_m *is termed the inverse GGDOP.*

Based on the above definition, the average localization error using m range estimates is given by

$$\Omega_{\mathbf{x},CRLB}(m) = \frac{\mathcal{G}_m}{\gamma_m} = \frac{1}{\gamma_m \Gamma_m}.$$

From the above relation, for a given set of range estimate variances, as the GGDOP $\mathcal{G}_m$ increases, the localization error increases. Under the condition that the range estimate variances are equal, i.e., $\sigma_i^2 = \sigma^2$, $i = 1, 2, \cdots, m$,

$$\mathcal{G}_m = \frac{m^2}{\sum_{i=1}^{m} \sum_{j=1, j>i}^{m} \sin^2(\alpha_i - \alpha_j)}.$$

This implies that the GGDOP is dependent only on the number and geometry of anchor nodes that provide range estimates when the range estimate variances are identical. Note that this is a generalized version of the classical GDOP definition [9], which constitutes the case $\sigma_i^2 = \sigma^2,\ \forall i$. For a given value of γ_m, as the GGDOP $\mathcal{G}_m$ increases, the localization error increases. It must be noted that $\mathcal{G}_m$ is not purely a function of the geometry of unlocalized nodes, but also depends on the range estimate variances. However, scaling all the range estimate variances by a common factor while maintaining the same relative orientations does not alter the value of $\mathcal{G}_m$. The following theorem shows that the value of $\mathcal{G}_m$ is lower bounded. The proof is provided in Appendix 3B.

Theorem 3 (Bounds on GGDOP). *The minimum and maximum possible values of* Γ_m *are* 0 *and* $\frac{1}{4}$ *respectively:*

$$0 \leq \Gamma_m = \frac{\psi_m}{\gamma_m^2} \leq \frac{1}{4} \Rightarrow 4 \leq \mathcal{G}_m = \frac{\gamma_m^2}{\psi_m} < \infty.$$

The maximum value of Γ_m *occurs when the node to be localized lies at the "center-of-mass" of a modified system of anchor nodes, where "mass" is analogous to the reciprocal of the range estimate variance. The orientation of the anchor nodes in this hypothetical system is obtained by doubling the angles in the original system. Explicitly,* $\mathcal{G}_m = 4$ *if and only if*

$$\sum_{i=1}^{m} \frac{\cos 2\alpha_i}{\sigma_i^2} = 0,\ \sum_{i=1}^{m} \frac{\sin 2\alpha_i}{\sigma_i^2} = 0.$$

Corollary (Localization Error with Optimal Geometry). *The localization error in the optimal geometric configuration* $\mathcal{G}_m = 4$ *is given by*

$$\Omega^*_{\mathbf{x},CRLB} = \frac{4}{\gamma_m} = \frac{4}{\sum_{i=1}^{m} \frac{1}{\sigma_i^2}}. \tag{3.7}$$

The above equation indicates that in the optimal geometrical scenario, as the range estimate variances decrease, the average localization error also decreases. The following theorem quantifies the impact of range estimate variances on the average localization error for an arbitrary geometry of anchor nodes. The proof is given in Appendix 3C.

Theorem 4 (Dependence on σ_i^2). *Suppose we are given an initial geometric configuration of anchor nodes* $\{\alpha_i\}$, $i = 1, 2, \cdots, m$ *with the corresponding range estimate variances* $\{\sigma_i^2\}$. *The localization error is then given by (3.5).*

The decrease in any *of the variances results in the reduction of the localization error. Specifically, for a given* i, *if* $\sigma_i'^2 = a\sigma_i^2$, $0 < a \leq 1$, *the new localization error* $\Omega'_{\mathbf{x},CRLB}$ *satisfies*

$$\Omega'_{\mathbf{x},CRLB} = \frac{\sum_{k=1,k\neq i}^{m} \frac{1}{\sigma_k^2} + \frac{1}{a\sigma_i^2}}{\sum_{k=1,k\neq i}^{m} \sum_{l=1,l>k,l\neq i}^{m} \frac{\sin^2(\alpha_k-\alpha_l)}{\sigma_k^2\sigma_l^2} + \frac{1}{a\sigma_i^2}\sum_{k=1,k\neq i}^{m} \frac{\sin^2(\alpha_k-\alpha_i)}{\sigma_k^2}} \leq \Omega_{\mathbf{x},CRLB}. \tag{3.8}$$

The above theorem states that a reduction of any range estimate variance results in the decrease of the localization error. From the perspective of PoLoNet design, this implies that repeated range estimation followed by averaging can also improve localization accuracy.

Theorem 5 (Dependence on m). *Suppose we have an initial geometric configuration of anchor nodes $\{\alpha_i\}$ with range estimate variances $\{\sigma_i^2\}$, $i = 1, 2, \cdots, m$. Then, the localization error is a function of m given by (3.5). The introduction of a new node with orientation and variance $\left(\alpha_{m+1}, \sigma_{m+1}^2\right)$* always *results in the reduction of the localization error. Specifically,*

$$\Omega_{\mathbf{x},CRLB}(m+1) = \frac{\sigma_{m+1}^2 \gamma_m + 1}{\sigma_{m+1}^2 \psi_m + \frac{\gamma_m}{2} - \frac{\sqrt{\gamma_m^2 - 4\gamma_m}}{2} \cos\left(2\alpha_{m+1} - 2\nu\right)} \leq \Omega_{\mathbf{x},CRLB}(m), \tag{3.9}$$

where

$$\nu = \frac{1}{2} \arctan\left(\frac{\sum_{i=1}^{m} \frac{\sin(2\alpha_i)}{\sigma_i^2}}{\sum_{i=1}^{m} \frac{\cos(2\alpha_i)}{\sigma_i^2}} \right). \tag{3.10}$$

The reduction in the average localization error as the number of range estimates is increased from m to $(m+1)$ can be shown to be

$$\left[\frac{\left(\sqrt{1 - 4\Gamma_m} - \cos\left(2\alpha_{m+1} - 2\nu\right)\right)^2 + \sin^2\left(2\alpha_{m+1} - 2\nu\right)}{2\Gamma_m \left(\sigma_{m+1}^2 \psi_m + \frac{\gamma_m}{2} - \frac{\sqrt{\gamma_m^2 - 4\psi_m} \cos\left(2\alpha_{m+1} - 2\nu\right)}{2} \right)} \right]. \tag{3.11}$$

Corollary (Optimum Orientation of Introduced Node). *Irrespective of the range variance of the introduced anchor node, the angles $\alpha_{m+1} = \alpha^*$ and $\alpha_{m+1} = \alpha^\dagger$ which respectively minimize and maximize $\Omega_{\mathbf{x},CRLB}(m+1)$ for a fixed initial configuration with m nodes, are given by*

$$\alpha^\star = \nu \pm \frac{\pi}{2}, \ \alpha^\dagger = \nu,$$

where ν is defined in (3.10).

Corollary (Repeated Measurements). *In terms of the reduction of localization error, the special case where $\alpha_{m+1} = \alpha_k$, $\sigma_{m+1} = \sigma_k$ where $k \in \{1, 2, \cdots, m\}$, is equivalent to a repeated range measurement followed by averaging, which reduces the variance of the range estimate by a factor of 2 (i.e., $a = \frac{1}{2}$ in Theorem 4).*

The proof of the above results is provided in Appendix 3D. From the above theorem, increasing the number of available range estimates always decreases the localization error of a node. We summarize the above theorems on the characteristics of the CRLB below:

1. The average localization error depends not only on the accuracy and number of the available range estimates, but also on the relative orientations of anchors providing those range estimates.

2. For a given geometry, the decrease in the variances of any subset of range estimates results in the decrease of the average localization error. Therefore, repeated range estimation followed by averaging results in the reduction of the average localization error.

3. Increasing the number of available range estimates always improves performance, irrespective of range variance and geometry. However, the improvement can be maximized with respect to geometry by carefully selecting the orientation of the introduced node.

It is interesting to note the CRLB with exponentially distributed range estimates (see Appendix 3E) is similar in structure to the CRLB for unbiased Gaussian range estimates discussed above. This suggests that the above properties, albeit in slightly modified form, are also observed for the case of exponentially distributed range estimates.

3.3.1 CRLB and Error Ellipses

As discussed previously, the CRLB bounds the eigenvalues of the covariance matrix $\mathbf{C}_{\hat{\mathbf{x}}}$ of the location estimate $\hat{\mathbf{x}}$:

$$\mathbf{C}_{\hat{\mathbf{x}}} \geq \mathbf{I}^{-1} \Rightarrow \mathbf{C}_{\hat{\mathbf{x}}} - \mathbf{I}^{-1} \geq \mathbf{0},$$

where $\mathbf{I}$ is the Fisher information matrix defined in (3.2). In the quantification of the attainable localization accuracy through the average localization error $\Omega_{\mathbf{x},CRLB}$, we ignored the off-diagonal elements of $\mathbf{C}_{\hat{\mathbf{x}}}$. However, the elements represent the correlation between the x and y coordinates of the location estimate, and are non-zero in general. If the location-estimate $\hat{\mathbf{x}}$ is assumed to be an unbiased Gaussian estimate of the true location $\mathbf{x}$, then the probability density function of the error in the location estimate $\mathbf{e} = \mathbf{x} - \hat{\mathbf{x}}$ is given by

$$f_{\mathbf{e}}(\mathbf{e}) = \frac{1}{(2\pi)^2\sqrt{\det \mathbf{C}_{\hat{\mathbf{x}}}}} \exp\left(-\frac{\mathbf{e}^T \mathbf{C}_{\hat{\mathbf{x}}}{}^{-1}\mathbf{e}}{2}\right). \tag{3.12}$$

This implies that equi-probable error contours are given by the following condition on the quadratic form:

$$\mathbf{e}^T \mathbf{C}_{\hat{\mathbf{x}}}{}^{-1}\mathbf{e} = k,$$

where k determines the probability of observing the error $\mathbf{e}$. This above relation corresponds to an ellipse in the two-dimensional case, and an ellipsoid in the three-dimensional case. Therefore, a useful way of visualizing the CRLB for two-dimensional location estimation is through the use of error ellipses. The larger the error ellipse, the larger the localization error. Further, the orientation of the error ellipse illustrates the impact of geometry on the accuracy of location-estimates. The value of k is selected such that [46] the probability that

the location-estimate lies inside the error-ellipse is[1] p:

$$k = -2\log(1-p)$$

If the (positive) eigenvalues of $\mathbf{C}_{\hat{\mathbf{x}}}$ are λ_1 and λ_2, with $\lambda_1 \geq \lambda_2$, the eigenvalues of $\mathbf{C}_{\hat{\mathbf{x}}}{}^{-1}$ are then $\frac{1}{\lambda_1}$ and $\frac{1}{\lambda_2}$, with unchanged eigenvectors. The semi-major and semi-minor axes of the error ellipse are given by eigenvectors corresponding to eigenvalues λ_2 and λ_1 respectively. The lengths of the semi-major and semi-minor axes [46] are given by $\sqrt{k\lambda_1}$ and $\sqrt{k\lambda_2}$ respectively. This implies that decreasing the eigenvalues of $\mathbf{C}_{\mathbf{x}}$ decreases the localization error, since

$$\lambda_1 + \lambda_2 = \mathrm{Tr}\left(\mathbf{C}_{\mathbf{x}}\right) = \Omega_{\mathbf{x},CRLB}. \tag{3.13}$$

The area of the ellipse is given by

$$A_e = \pi k\sqrt{\lambda_1\lambda_2}.$$

It can be verified that the eigenvalues of the Fisher Information matrix $\mathbf{I}$ are given by

$$\lambda(\mathbf{I}) = \frac{\gamma_m}{2}\left(1 \pm \sqrt{1-4\Gamma_m}\right).$$

For the efficient estimator, the eigenvalues λ_1 and λ_2 of $\mathbf{C}_{\mathbf{x}}$ are given by

$$\lambda_1 = \frac{1}{\frac{\gamma_m}{2}\left(1-\sqrt{1-4\Gamma_m}\right)},\ \lambda_2 = \frac{1}{\frac{\gamma_m}{2}\left(1+\sqrt{1-4\Gamma_m}\right)}. \tag{3.14}$$

It is easy to verify that

$$\lambda_1 + \lambda_2 = \frac{1}{\gamma_m\Gamma_m} = \Omega_{\mathbf{x}},$$

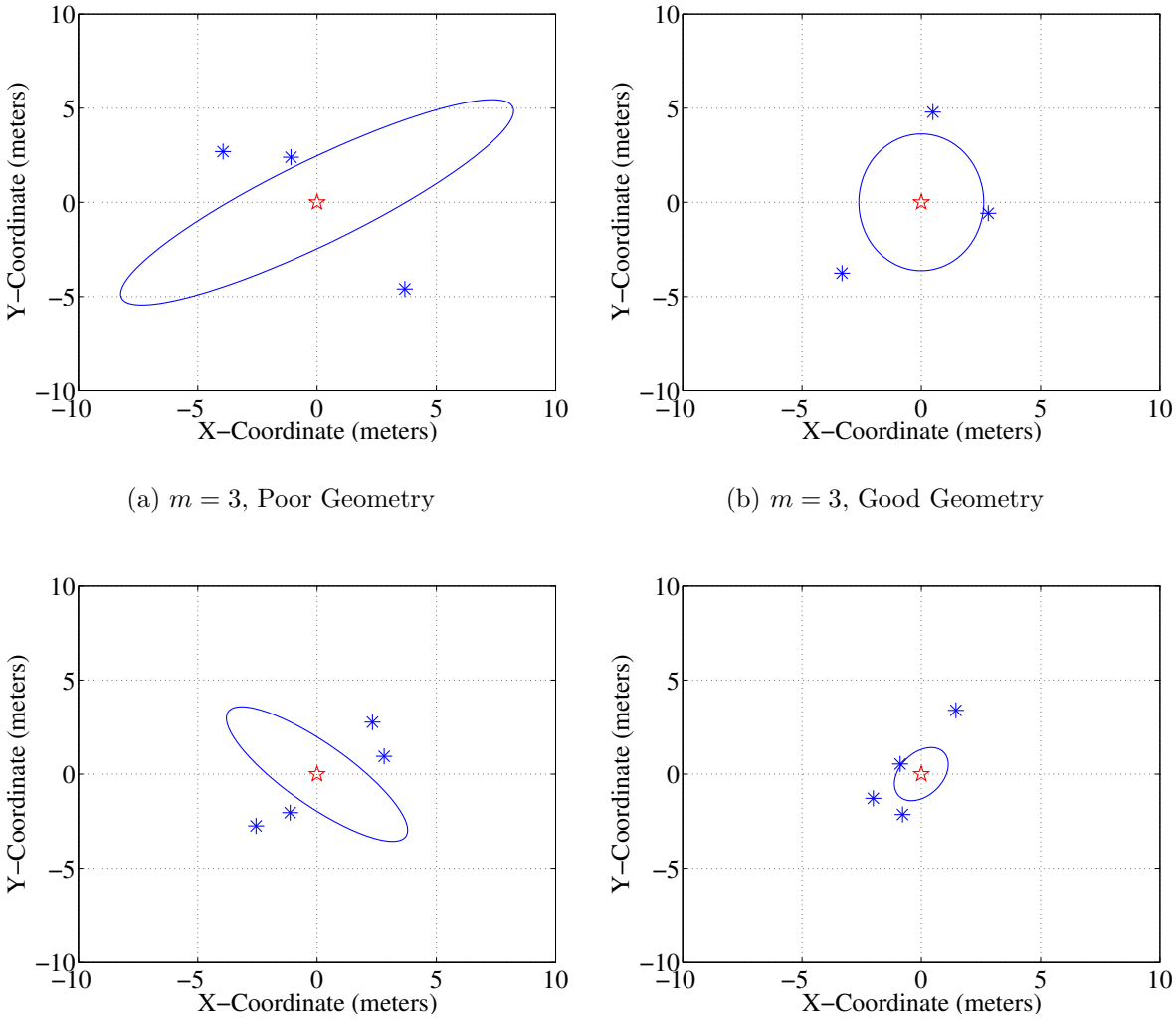
as predicted by (3.13). We see from (3.14) that if $\Gamma_m \to 0$ (worst-case geometry), $\lambda_1 \to \infty$ and thus the localization error tends to infinity. Similarly, if $\Gamma_m = \frac{1}{4}$ (best-case geometry), $\lambda_1 = \lambda_2 = \frac{2}{\gamma_m}$. This implies that the error ellipse for any value of k (or p) reduces to a circular region with area $A_e = \frac{4\pi k}{\gamma_m}$. Figure 3.2 illustrates the impact of anchor geometry and the number of range estimates on the CRLB error ellipses.

3.4 The Least-Squares (LS) estimator

The CRLB in terms of the average localization error given by (3.5) provides a benchmark for evaluating the performance of practical location estimators, but does not explicitly describe

[1]This value of k arises from a normalized, rotated ellipse transformed to polar coordinates:

$$p = \int_0^{\sqrt{k}}\int_0^{2\pi}\frac{1}{2\pi}\exp\left(-\frac{r^2}{2}\right)rdrd\theta = 1-\exp\left(-\frac{k}{2}\right).$$



(a) $m = 3$, Poor Geometry

(b) $m = 3$, Good Geometry

(c) $m = 4$, Poor Geometry

(d) $m = 4$, Good Geometry

Figure 3.2: Examples of CRLB Error ellipses with $p = 0.95$: The anchor nodes (asterisks) are assumed to be uniformly distributed over a 10×10 meter^2 area, and the unlocalized node is located at the origin. The range estimates are assumed to be unbiased and Gaussian distributed with variance $\sigma_i^2 = 0.1R_i^2$. We see that as m increases, the size of the error ellipses decrease.

the estimator that achieves it. The Least-Squares (LS) estimation approach is known to be suitable if a statistical model for the available data (range estimates in this case) is not known [34]. Additionally, it was shown in [20] that the LS estimator is more robust to range bias errors than other estimators.

If the coordinates of m anchor nodes $\mathbf{x}_i$, $(i = 1, 2, \cdots, m)$ and the corresponding range estimates r_i are known, then $\hat{\mathbf{x}}$ can be computed based on the LS formulation as follows. We have m equations of the form

$$\|\mathbf{x} - \mathbf{x}_i\|^2 = r_i^2, \quad i = 1, 2, \cdots, m.$$

The above equations can be simplified to the form:

$$\|\mathbf{x}\|^2 + \|\mathbf{x}_i\|^2 - 2\mathbf{x}_i^T\mathbf{x} = r_i^2, \quad i = 1, 2, \cdots, m.$$

Taking the difference between pairs of the above m equations, this system reduces to a set of $\binom{m}{2}$ linear equations:

$$\begin{aligned} 2(\mathbf{x}_i - \mathbf{x}_j)^T\mathbf{x} &= (R_{i0}^2 - R_{j0}^2) - (r_i^2 - r_j^2), \\ i &= 1, 2, \cdots, m-1, \quad j > i, \end{aligned}$$

where $R_{i0} = \|\mathbf{x}_i\|$. It is important to point out that we must have at least $m = 3$ in order to obtain two linear equations that can be solved simultaneously. Therefore, for two-dimensional location estimation, the LS estimator requires $m \geq 3$. The above set of linear equations is of the form

$$\mathbf{A}\mathbf{x} = \mathbf{b}, \tag{3.15}$$

where $\mathbf{A}$ is a $\binom{m}{2} \times 2$ matrix whose rows are from the set[2] $\{2(\mathbf{x}_i - \mathbf{x}_j)^T\}$, $(i, j = 1, 2, \cdots, m,\ j > i)$ and $\mathbf{b}$ is a column vector of length $\binom{m}{2}$ whose components are from $\{(R_{i0}^2 - R_{j0}^2) - (r_i^2 - r_j^2)\}$. The LS estimate of the node's location is given by

$$\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}. \tag{3.16}$$

The average localization error for the LS estimator is defined as

$$\Omega_{\mathbf{x},LS} = E\left\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\right\}.$$

It can seen from the equations that the model for the range estimates cannot be expressed as a linear function of $\mathbf{x}$:

$$\mathbf{r} \neq \mathbf{H}\mathbf{x} + \mathbf{w}.$$

Therefore, the LS estimator is not the minimum variance unbiased estimator (MVUE) [34]. This is also seen in Figure 3.3, where the performance of the LS estimator is compared with

[2]Note that $\mathbf{a} = \{\alpha\}$ is used to represent a generic component of the vector (or matrix) $\mathbf{a}$ without reference to the specific ordering of components (or rows).

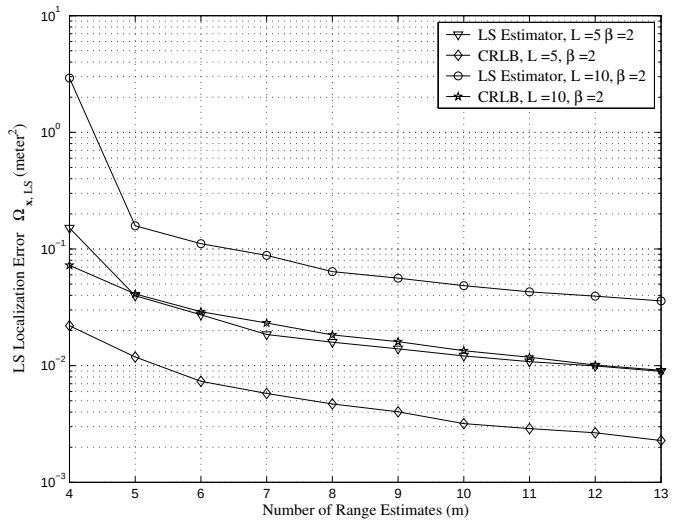


Figure 3.3: Comparison of the LS estimator with the CRLB for two-dimensional location estimation given m unbiased Gaussian range estimates.

the CRLB in terms of the average localization error given unbiased Gaussian (noisy) range estimates. The details of the simulation are as follows: m anchor nodes whose locations are known exactly are dispersed randomly over an $L \times L$ area. The range estimates r_i are assumed to be unbiased Gaussian random variables with variance $\sigma_i^2 = K_E R_i^\beta$. The values of K_E and β used for this simulation were 0.01 and 2 respectively. As expected, the LS estimator performs worse than the CRLB in terms of the localization error. We also make another important observation: as shown with the CRLB in Theorem 5, the LS localization error decreases with increasing m.

Figure 3.4 shows the sensitivity of the LS estimate to the geometry of nodes, when m is increased. In this simulation, the locations of the first $(m-1)$ anchor nodes was fixed, and the orientation α_m of the mth anchor node was varied between $[0, 2\pi)$. We observe that increasing the number of available range estimates not only decreases the LS localization error, but also considerably reduces its sensitivity to geometry.

Figure 3.5 shows the average localization error of the LS location estimator when the range estimates *and* the anchor location estimates are noisy. We see that even with noisy location estimates of the anchor nodes $\hat{\mathbf{x}}_i \sim \mathcal{N}(\mathbf{x}_i, \sigma_P^2 \mathbf{I})$, increasing the number of available range estimates decreases the localization error. It must be noted that this is a realistic scenario in ad hoc UWB PoLoNets, since the locations of unlocalized reference nodes are computed using noisy range estimates from fixed anchors and localized reference nodes. In the following subsections, we show that for small range measurement noise in the presence of perfect anchor location estimates, the localization error for the LS estimate can be approximated by a Gaussian random variable.

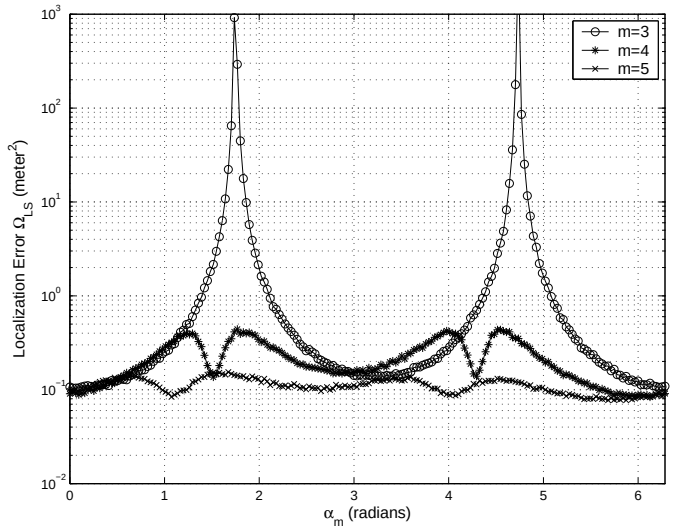


Figure 3.4: Impact of geometry on the LS location estimate. In this simulation, $K_E = 0.01$, $L = 10$ meters.

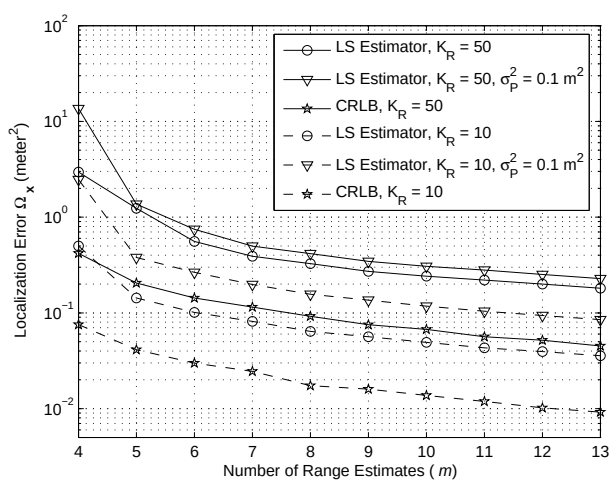


Figure 3.5: Average localization error versus the number of range estimates available at a mobile node using an LS estimator with (a) exact reference node coordinates $\hat{\mathbf{x}}_i = \mathbf{x}_i$ and noisy range estimates $r_i \sim \mathcal{N}(R_i, \sigma_i^2)$, $\sigma_i^2 = \frac{K_R R_i^\beta}{K_P P_t}$, $\beta = 2$, $K_P = 1 \times 10^6$, $P_t = 1$ mW, and (b), noisy reference node coordinates $\hat{\mathbf{x}}_i \sim \mathcal{N}(\mathbf{x}_i, \sigma_P^2 \mathbf{I})$ and noisy range estimates $r_i \sim \mathcal{N}(R_i, \sigma_i^2)$.

3.4.1 Distribution of the Localization Error

From (2.25), the noisy range measurements are modeled as

$$r_i = R_i + n_i,\ i = 1, 2, \cdots, m,$$

where $n_i \sim \mathcal{N}(0, \sigma_i^2)$. Assuming the range measurement noise is small when compared to the actual inter-node distances, $n_i^2 \ll R_i,\ i = 1, 2, \cdots, m$, we can rewrite the expression for $\mathbf{b}$ in (3.15), where $\forall\ i, j = 1, 2, \cdots, m,\ i < j$:

$$\begin{aligned}\tilde{\mathbf{b}} &= \{(R_{i0}^2 - R_{j0}^2) - (r_i^2 - r_j^2)\} = \{(R_{i0}^2 - R_{j0}^2) - \left[(R_i + n_i)^2 - (R_j + n_j)^2\right]\} \\ &\approx \{\left[(R_{i0}^2 - R_{j0}^2) + (R_i^2 - R_j^2)\right] - 2\left[R_i n_i - R_j n_j\right]\} = \mathbf{b} - 2\{R_i n_i - R_j n_j\} = \mathbf{b} - \epsilon_{\mathbf{b}}.\end{aligned}$$

where

$$\epsilon_{\mathbf{b}} = 2\{R_i n_i - R_j n_j\},\ i, j = 1, 2, \cdots, m,\ i < j.$$

Assuming that the range errors n_i, n_j are independent, it can be shown that the components of the vector $\epsilon_{\mathbf{b}}$ are also Gaussian: $2\left(R_i n_i - R_j n_j\right) \sim \mathcal{N}\left(0, 4\left(R_i^2\sigma_i^2 + R_j^2\sigma_j^2\right)\right)$. The estimate of the node's location is given by

$$\hat{\mathbf{x}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\hat{\mathbf{b}} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\left(\mathbf{b}^0 - \epsilon_{\mathbf{b}}\right) \Rightarrow \mathbf{x} - \hat{\mathbf{x}} = \left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\epsilon_{\mathbf{b}}.$$

From the above equation, we see that the location estimate error is $\mathbf{x} - \hat{\mathbf{x}}$ is a linear combination of Gaussian random variables, and is therefore also Gaussian distributed. The localization error is given by

$$\begin{aligned}\Omega_{\mathbf{x},LS} &= E\left[(\mathbf{x} - \hat{\mathbf{x}})^T(\mathbf{x} - \hat{\mathbf{x}})\right] = \mathrm{Tr}\left\{E\left[(\mathbf{x} - \hat{\mathbf{x}})(\mathbf{x} - \hat{\mathbf{x}})^T\right]\right\} \\ &= \mathrm{Tr}\left\{E\left[\left(\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\epsilon_{\mathbf{b}}\right)\left(\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\epsilon_{\mathbf{b}}\right)^T\right]\right\} \\ &= \mathrm{Tr}\left\{\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\mathbf{A}^T\mathbf{C}_{\epsilon_{\mathbf{b}}}\mathbf{A}\left(\mathbf{A}^T\mathbf{A}\right)^{-1}\right\},\end{aligned} \tag{3.17}$$

where $\mathbf{C}_{\epsilon_{\mathbf{b}}}$ is the covariance matrix of $\epsilon_{\mathbf{b}}$. From the above expression, we see that extracting a closed-form expression of the localization error in terms of the anchor coordinates and the range estimate variances is a tedious task, even for a small value of m. The above expression is much simplified in the efficient weighted LS case [34], when the covariance matrix of the range estimates is known, as will be seen in the following section.

3.5 Reliability and Weighted Least Squares

Based on the model of TOA-based range estimation discussed in Chapter 2, the variance of a range estimate was shown to be dependent of the received SNR, and therefore, on

the distance between unlocalized nodes and anchors. Consequently, range estimates from near-by anchors are likely to be more accurate than distant anchors. As a consequence, it is reasonable to weight range estimates from nearby anchors higher than range estimates from distant anchors. In the following section, we derive the Weighted Least-Squares (WLS) estimate for a nodes location given unbiased range estimates from anchors, and show that the optimal weighting matrix cannot be implemented in practice as it depends on the true distances between nodes, which are unknown. We provide a heuristic modification of the optimal weighting matrix which still provides sizable gains in terms of localization accuracy over the LS method discussed in Section 3.4.

3.5.1 Simple Weighted Least Squares

Given m range estimates r_i, $i = 1, 2, \cdots, m$, and the corresponding anchor locations, $\mathbf{x}_i$, $i = 1, 2, \cdots, m$, we can write

$$(x - x_i)^2 + (y - y_i)^2 = R_i^2, \quad i = 1, 2, \cdots, m. \tag{3.18}$$

where $R_i = \|\mathbf{x} - \mathbf{x}_i\|$ is the true distance between the unlocalized node and the ith anchor. Suppose we take the difference between the ith equation ($i > 1$) above, and the first ($i = 1$), we obtain

$$(x_1 - x_i)x + (y_1 - y_i)y = \frac{R_i^2 - R_1^2 - (x_i - x_1)^2 - (y_i - y_1)^2}{2}, \quad i = 2, 3, \cdots, m.$$

This can be expressed in matrix form as

$$\mathbf{A}_1\mathbf{x} = \mathbf{b}_1, \tag{3.19}$$

where, as usual, $\mathbf{x} = [x\ y]^T$, and

$$\mathbf{A}_1 = \begin{bmatrix} (x_1 - x_2) & (y_1 - y_2) \\ (x_1 - x_3) & (y_1 - y_3) \\ \vdots & \vdots \\ (x_1 - x_m) & (y_1 - y_m) \end{bmatrix}_{(m-1)\times 2}, \quad \mathbf{b}_1 = \begin{bmatrix} \frac{R_2^2 - R_1^2 - (x_2 - x_1)^2 - (y_2 - y_1)^2}{2} \\ \frac{R_3^2 - R_1^2 - (x_3 - x_1)^2 - (y_3 - y_1)^2}{2} \\ \vdots \\ \frac{R_m^2 - R_1^2 - (x_m - x_1)^2 - (y_m - y_1)^2}{2} \end{bmatrix}_{(m-1)\times 1}.$$

The subscripts of matrices above indicate that the range differences were computed with respect to range estimate from anchor $i = 1$. It must be emphasized that this formulation is different from the LS formulation in Section 3.4, where the difference between *all* pairs of range estimates was used to solve for the node's location. Although a similar formulation can be adopted for a WLS method, it is less mathematically tractable than the method described here.

As the true range estimates R_i are unknown, the available noisy version of $\mathbf{b}_1$ is given by

$$\tilde{\mathbf{b}}_1 = \begin{bmatrix} \frac{r_2^2 - r_1^2 - (x_2-x_1)^2 - (y_2-y_1)^2}{2} \\ \frac{r_3^2 - r_1^2 - (x_3-x_1)^2 - (y_3-y_1)^2}{2} \\ \vdots \\ \frac{r_m^2 - r_1^2 - (x_m-x_1)^2 - (y_m-y_1)^2}{2} \end{bmatrix}_{(m-1)\times 1}.$$

Since $r_i = R_i + n_i$, we can write

$$r_i^2 - r_1^2 = (R_i + n_i)^2 - (R_1 + n_1)^2 \approx R_i^2 - R_1^2 + 2R_i n_i - 2R_1 n_1,$$

when $n_i \ll R_i$, $i = 1, 2, \cdots, m$. Under these conditions, we can approximate $\tilde{\mathbf{b}}_1$ using

$$\tilde{\mathbf{b}}_1 \approx \mathbf{b}_1 + \mathbf{w}_1 \tag{3.20}$$

where

$$\mathbf{w}_1 = \begin{bmatrix} R_2 n_2 - R_1 n_1 \\ R_3 n_3 - R_1 n_1 \\ \vdots \\ R_m n_m - R_1 n_1 \end{bmatrix}_{(m-1)\times 1} = \underbrace{\begin{bmatrix} -R_1 & R_2 & 0 & \cdots & 0 \\ -R_1 & 0 & R_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -R_1 & 0 & 0 & \cdots & R_m \end{bmatrix}_{(m-1)\times m}}_{\mathbf{D}} \underbrace{\begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_m \end{bmatrix}_{m\times 1}}_{\mathbf{n}}.$$

As $\mathbf{n}$ is a zero-mean vector of random variables, $\mathbf{w}_1$ is also zero-mean. Assuming that the measurement noise terms in different range estimates are uncorrelated, the covariance matrix of $\mathbf{w}_1$ is given by

$$\mathbf{C}_{\mathbf{w_1}} = \mathbf{D}\mathbf{\Lambda_n}\mathbf{D}^T, \tag{3.21}$$

where $\mathbf{\Lambda_n}$ is the diagonal matrix containing noise variances:

$$\mathbf{\Lambda_n} = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_m^2 \end{bmatrix}_{m\times m}.$$

From (3.19) and (3.20), we can write

$$\tilde{\mathbf{b}}_1 \approx \mathbf{A}_1\mathbf{x} + \mathbf{w}_1. \tag{3.22}$$

For this model, the Best Linear Unbiased Estimator (BLUE) [34] is well-known:

$$\hat{\mathbf{x}} = \left(\mathbf{A}_1^T\mathbf{W}\mathbf{A}_1\right)^{-1}\mathbf{A}_1^T\mathbf{W}\tilde{\mathbf{b}}_1, \tag{3.23}$$

where the optimal weighting matrix $\mathbf{W}$ is the inverse of the covariance matrix $\mathbf{C}_{\mathbf{w_1}}$:

$$\mathbf{W} = \mathbf{C}_{\mathbf{w_1}}{}^{-1}.$$

In essence, the weighting matrix weights range estimates from nearby anchors higher than those from farther away anchors since the range estimation accuracy is inversely proportional to distance. From (3.21), we see the stumbling block in using the above WLS estimator. The covariance matrix $\mathbf{C_{w_1}}$ depends on the *true* ranges R_i, $i = 1, 2, \cdots, m$ through the matrices $\mathbf{D}$ and $\mathbf{\Lambda_n}$ (as the range estimate variances are determined by the true distances R_i), which are unknown. Hence, the optimal weighting matrix cannot be computed in this manner.

A heuristic adaptation of the WLS method above would be to replace the true ranges R_i in the definition of $\mathbf{W}$, with the range estimates r_i. This implies that the new weighting matrix $\tilde{\mathbf{W}}$ is given by

$$\tilde{\mathbf{W}} = \tilde{\mathbf{C}}_{\mathbf{w_1}}^{-1} = \left(\tilde{\mathbf{D}}\tilde{\mathbf{\Lambda}}_{\mathbf{n}}\tilde{\mathbf{D}}^T\right)^{-1},$$

where

$$\tilde{\mathbf{D}} = \begin{bmatrix} -r_1 & r_2 & 0 & \cdots & 0 \\ -r_1 & 0 & r_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -r_1 & 0 & 0 & \cdots & r_m \end{bmatrix}_{(m-1)\times m}, \quad \mathbf{\Lambda_n} = \begin{bmatrix} \hat{\sigma}_1^2 & 0 & \cdots & 0 \\ 0 & \hat{\sigma}_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \hat{\sigma}_m^2 \end{bmatrix}_{m\times m}. \tag{3.24}$$

In the above equation, since the range estimate variances σ_i^2 depend on R_i, $i = 1, 2, \cdots, m$, we can use $\hat{\sigma}_i^2 = Kr_i^{\hat{\beta}}$, where $\hat{\beta}$ is an estimate of the path loss exponent. If such an estimate is not available, we can simply use $\hat{\sigma}_i^2 = Kr_i^2$. Note from (3.23) that the constant K does not affect the performance of the estimator.

Figure 3.6 compares the performance of the heuristic WLS method discussed above with the LS estimator formulated in section 3.4, and the CRLB. We see that the WLS method outperforms the LS estimator and is closer to the CRLB. Further, relative to the LS estimator, the gains in localization accuracy increase as m increases.

Figure 3.7 compares the performance of the heuristic WLS method discussed above and "ideal" WLS, assuming the optimal weighting matrix is available. We observe that (i) for small values of K_E, the heuristic and ideal WLS are comparable, but (ii) as K_E increases, the performance of the heuristic WLS rapidly degrades relative to the ideal WLS solution, as the weighting matrix is no longer close to the optimal weighting matrix[3]. Therefore, for small range estimate variances, the performance of the heuristic WLS method can be quantified using [34]

$$\Omega_{\mathbf{x},WLS} \approx \mathrm{Tr}\left(\left(\mathbf{A}_1^T\mathbf{C}_{\mathbf{w}}^{-1}\mathbf{A}_1\right)^{-1}\right). \tag{3.25}$$

Although the above approximation for the WLS localization error is much simpler than the corresponding expression for the LS estimator given in (3.17), this does not directly associate the localization error with the distributions of the range measurement error and

[3]It has been observed via simulations that as the range estimate variances increase, the matrix $\mathbf{A}_1^T\mathbf{W}\mathbf{A}_1$ tends to become ill-conditioned with greater probability.

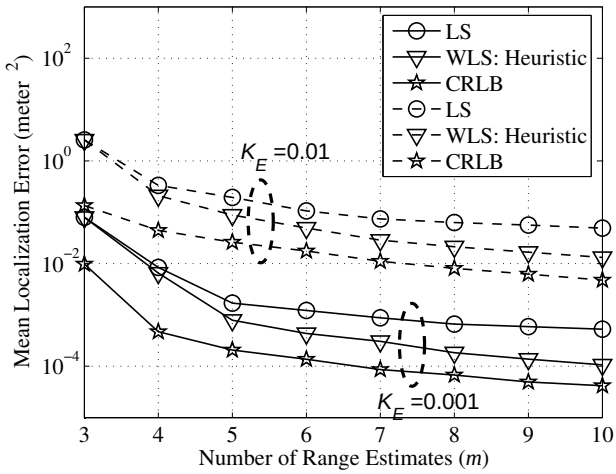


Figure 3.6: Comparison of the heuristic WLS method with the previously discussed LS method and the CRLB for different values of m and K_E. The path loss exponent was assumed to be $\beta = 2$. The plots depict the localization error averaged over a large number of realizations of range measurement noise and anchor locations.

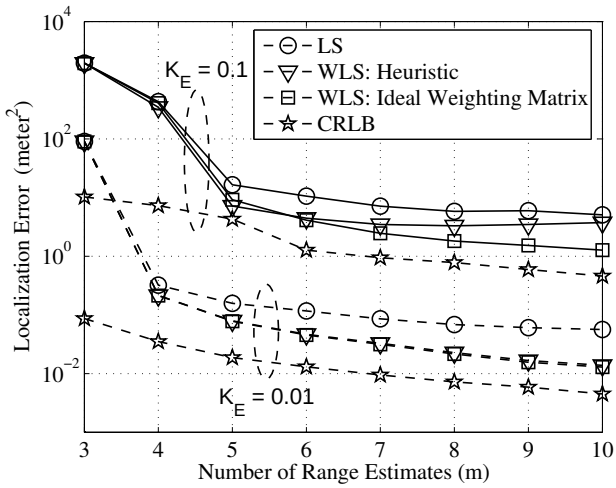


Figure 3.7: Comparison of the heuristic WLS with "ideal" WLS: we observe that (i) for small values of K_E, the heuristic and ideal WLS are comparable, but (ii) as K_E increases, the performance of the heuristic WLS rapidly degrades relative to the ideal WLS solution.

anchor locations. In conclusion, for arbitrary anchor locations and range error distributions, the CRLB remains the most tractable means of quantifying the localization error.

Robust localization through selection of a subset of range estimates can be viewed as a special case of the WLS estimator discussed above. Further enhancements of the above WLS methods are possible. In the case where the estimates of anchor locations are noisy, we can formulate a weighted LS solution where range information from anchors with more reliable location estimates are weighted more heavily than unreliable ones. Each anchor and its corresponding range estimate are associated with a reliability metric (e.g., the reliability of a location estimate could be assigned based on the geometrical conditions under which it was computed), and the weights are computed using the reliability estimates.

The LS and WLS location-estimators discussed above provide "one-shot" estimates of a node's location. Using these as initial solutions, we can obtain more accurate location estimates through iterative refinement. In the following section, we quantify the gains in localization accuracy provided by such iterative refinement of location estimates.

3.6 Iterative Refinement of Location-Estimates

The problem of two-dimensional location estimation can be cast as the weighted minimization of the distance error residuals: Given $\{r_i\}$, $\{\mathbf{x}_i\}$, $i = 1, 2, \cdots, m$, find $\mathbf{x} = [x\ y]^T$ that minimizes the cost function:

$$\epsilon(\mathbf{x}) = \sum_{i=1}^{m} w_i (r_i - R_i(\mathbf{x}))^2 = \sum_{i=1}^{m} w_i \left(r_i - \sqrt{(x - x_i)^2 + (y - y_i)^2} \right)^2 . \tag{3.26}$$

We define the vector $\mathbf{e}(\mathbf{x})$ as

$$\mathbf{e}(\mathbf{x}) = [r_1 - R_1(\mathbf{x}),\ r_2 - R_2(\mathbf{x}),\ \cdots,\ r_m - R_m(\mathbf{x})]^T . \tag{3.27}$$

Then the cost function can be re-written as

$$\epsilon(\mathbf{x}) = \mathbf{e}(\mathbf{x})^T \mathbf{W} \mathbf{e}(\mathbf{x}), \tag{3.28}$$

where $\mathbf{W} = \mathrm{diag}\,\{w_1, w_2, \cdots, w_m\}$.

This is a nonlinear estimation problem and a closed-form solution does not exist [47]. However, a variety of optimization algorithms are available for solving this type of nonlinear estimation. For instance, the quasi-Newton DFP (Davidon-Fletcher-Powell) algorithm has been employed for locating valuable assets [48] and the Levenberg-Marquardt method was studied in [47]. Both unconstrained and constrained minimization algorithms [3] can be employed to achieve optimal position estimation.

3.6.1 Gauss-Newton Methods

The basis for Gauss-Newton methods lies in the Taylor-series expansion of the non-linear function $\epsilon(\mathbf{x})$:

$$\epsilon(\mathbf{x}_k + \mathbf{d}_k) \approx \epsilon(\mathbf{x}_k) + \mathbf{g}_k^T \mathbf{d}_k + \frac{\mathbf{d}_k^T \mathbf{H}(\mathbf{x}_k) \mathbf{d}_k}{2},$$

where $\mathbf{d}_k$ is the increment to be determined, and $\mathbf{g}_k$ and $\mathbf{H}(\mathbf{x}_k)$ are the gradient and Hessian of $\epsilon(\mathbf{x})$ at $\mathbf{x} = \mathbf{x_k}$ respectively. From the above approximation, the value of $\mathbf{d}_k$ that minimizes $\epsilon(\mathbf{x}_k + \mathbf{d}_k)$ is given by

$$\mathbf{H}(\mathbf{x}_k) \mathbf{d}_k = -\mathbf{g}_k,$$

using partial derivatives with respect to $\mathbf{d}_k$. The choice of $\mathbf{d}_k$ based on the above equation, which requires the inversion of the Hessian matrix, comprises Newton's method. In order to avoid computation of the Hessian $\mathbf{H}(\mathbf{x}_k)$ that uses second-order information, the Gauss-Newton method uses

$$\mathbf{J}_k^T \mathbf{J}_k \mathbf{d}_k = -\mathbf{J}_k^T \mathbf{e}(\mathbf{x}_k) \tag{3.29}$$

where $\mathbf{J}_k$ is the Jacobian of $\mathbf{e}(\mathbf{x})$ at $\mathbf{x} = \mathbf{x}_k$. If $\mathbf{J}_k$ is full rank, $\mathbf{d}_k$ can be determined by the linear least-squares solution:

$$\mathbf{d}_k = -\left(\mathbf{J}_k^T \mathbf{J}_k\right)^{-1} \mathbf{J}_k^T \mathbf{e}(\mathbf{x}_k) \tag{3.30}$$

The Gauss-Newton method has been observed to perform poorly when the second-order information in the Hessian is non-trivial. A modification of the Gauss-Newton method, called the Levenberg-Marquardt (LM) method, uses the following system of equations to solve for $\mathbf{d}_k$:

$$\left(\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \mathbf{I}\right)^{-1} \mathbf{d}_k = -\mathbf{J}_k^T \mathbf{e}(\mathbf{x}_k), \tag{3.31}$$

where λ_k is a non-negative scalar. The LM method uses a search direction that is a cross between the Gauss-Newton direction and the steepest descent, and has been to shown to be more robust than the Gauss-Newton method. The LM iteration is the default method used by the lsqnonlin script in Matlab's Optimization Toolbox. The LM method [47] has been shown to outperform the Quasi-Newton Davidon-Fletcher-Powell (DFP) iteration (which was used in [48]) in terms of localization accuracy.

One of the main issues with the use of iterative solutions is the choice of an initial solution. Since typical non-linear cost functions can have a very large number of local minima, the choice of the initial solutions plays a major role in determining the accuracy of the iterative solution. In general, an initial solution closer to the true solution results in a more accurate iterative solution.

Figure 3.8 compares the performance of the LS and WLS methods discussed in previous sections with the performance of an LM iterative solution that uses the LS and WLS estimates as initial solutions for iteration. We see that, at the cost of higher complexity, the LM iterative solution provides some gains in localization accuracy, especially for a small number

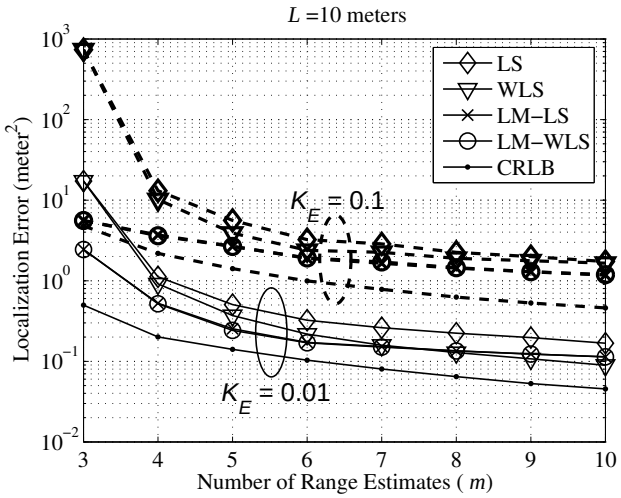


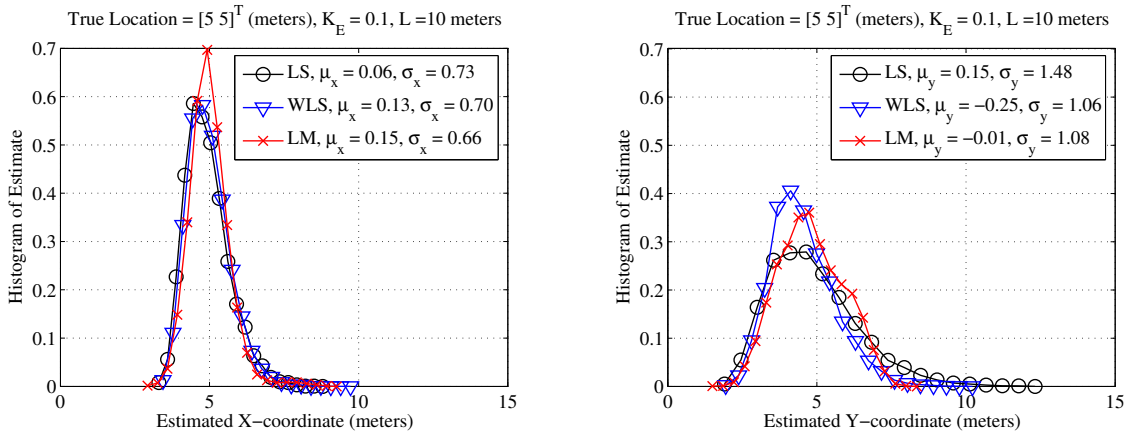
Figure 3.8: Comparison of iterative and non-iterative methods: the anchor nodes are assumed to randomly distributed over a 10×10 meter^2 area, and range estimate variances are generated according to $\sigma_i^2 = K_E R_i^2$. We observe that the use of iterative refinement improves the average localization accuracy, especially for small values of m. For small values of K_E we see that the WLS method outperforms the iterative methods for large values of m.

of range estimates. However, for small values of K_E, the heuristic WLS method outperforms the iterative methods for large values of m. We see that the use of WLS estimates as initial solutions for LM iteration results in marginally better location estimates than the use of LS estimates.

Figure 3.9 shows sample histograms of the x and y coordinates of location estimates generated using the LS, (heuristic) WLS and LM iterative methods. We observe that the estimated coordinates are biased, and cannot strictly be modeled as Gaussian random variables. We will see in Chapter 5 that this makes the modeling of the propagation of error problem with practical estimators more complicated.

3.7 Conclusions

In this chapter, we presented bounds on localization accuracy and compared them with the performance of the practical LS location estimation. Valuable analytical insights were gained using these bounds, with respect to the impact of node geometry, and the variance and number of range estimates on localization accuracy. These trends were also observed in LS estimation approaches, and the disparity between the practical estimators and the CRLB was quantified via simulation results. These insights are useful in the design of MAC protocols and Power-Control algorithms as will be shown in subsequent chapters. It was



(a) (b)

Figure 3.9: Histograms of estimated node coordinates for a single random location, with $m = 4$ unbiased range estimates with variances $\sigma_i^2 = K_E R_i^2$, $K_E = 0.1$. The anchors coordinates are assumed to be independently uniformly distributed over $[0L]$, $L = 10$ meters, with the unlocalized node located at $\mathbf{x} = [5\ 5]^T$ meters, and location estimates are generated for a large number of range measurement noise realizations. (a) Estimate of x-coordinate using LS, (heuristic) WLS and LM iterative with LS initial solution. (b) Estimate of y-coordinate using LS, (heuristic) WLS and LM iterative with LS initial solution. Also indicated are the means and standard deviations of the errors in the x and y coordinates.

further found that under the LS formulation, weighting and iterative refinement of location estimates can be used to considerably improve the average localization accuracy. Finally, due to a lack of tractability in terms of analysis of the LS estimators, the CRLB also serves as a measure of localization accuracy in subsequent analyses.

3.8 Appendix 3A: Proof of Equation (3.5)

The vector parameter to be estimated is

$$\theta = \mathbf{x} = [x \ y]^T$$

The given data are

$$\mathbf{r} = [r_1 \ r_2 \ \cdots \ r_m]^T$$

where $r_i, \ i = 1, 2, \cdots, m$ are the noisy estimates of the true ranges such that $n_i = r_i - \|\mathbf{x} - \mathbf{x}_i\| \sim \mathcal{N}(0, \sigma_i^2)$. We assume that each of the range measurements are independent. Therefore, the PDF

$$\begin{aligned} f(\mathbf{r};\theta) &= \prod_{i=1}^{m} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{(r_i - \|\mathbf{x} - \mathbf{x}_i\|)^2}{2\sigma_i^2}\right] \\ \Rightarrow \log f(\mathbf{r};\theta) &= -\log\left[(2\pi)^{\frac{m}{2}} \prod_{i=1}^{m} \sigma_i^2\right] - \sum_{i=1}^{m} \frac{(r_i - \|\mathbf{x} - \mathbf{x}_i\|)^2}{2\sigma_i^2} \end{aligned}$$

The Fisher information matrix is given by

$$\mathbf{I} = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}$$

where

$$\begin{aligned} I_{xx} &= -E\left[\frac{\partial^2}{\partial x^2} \log f(\mathbf{r};\theta)\right], \ I_{xy} = -E\left[\frac{\partial^2}{\partial x \partial y} \log f(\mathbf{r};\theta)\right] \\ I_{yx} &= -E\left[\frac{\partial^2}{\partial y \partial x} \log f(\mathbf{r};\theta)\right], \ I_{yy} = -E\left[\frac{\partial^2}{\partial y^2} \log f(\mathbf{r};\theta)\right] \end{aligned}$$

We obtain

$$\begin{aligned} \frac{\partial}{\partial x} \log f(\mathbf{r};\theta) &= -\frac{\partial}{\partial x} \sum_{i=1}^{m} \frac{(r_i - \|\mathbf{x} - \mathbf{x}_i\|)^2}{2\sigma_i^2} = -\sum_{i=1}^{m} \frac{1}{2\sigma_i^2} \frac{\partial}{\partial x} (r_i - \|\mathbf{x} - \mathbf{x}_i\|)^2 \\ &= -\sum_{i=1}^{m} \frac{1}{2\sigma_i^2} \frac{\partial}{\partial x} \left(r_i^2 - 2r_i\sqrt{(x - x_i)^2 + (y - y_i)^2} + (x - x_i)^2 + (y - y_i)^2\right) \\ &= \sum_{i=1}^{m} \frac{1}{2\sigma_i^2} \frac{\partial}{\partial x} \left(2r_i\sqrt{(x - x_i)^2 + (y - y_i)^2} - (x - x_i)^2\right) \\ &= \sum_{i=1}^{m} \frac{1}{\sigma_i^2} \left(r_i \frac{x - x_i}{\sqrt{(x - x_i)^2 + (y - y_i)^2}} - (x - x_i)\right) \end{aligned}$$

From the above equation, we obtain

$$\begin{aligned}
\frac{\partial^2}{\partial x^2}\log f(\mathbf{r};\theta) &= \frac{\partial}{\partial x}\sum_{i=1}^{m}\frac{1}{\sigma_i^2}\left(r_i\frac{x-x_i}{\sqrt{(x-x_i)^2+(y-y_i)^2}}-(x-x_i)\right)\\
&= \sum_{i=1}^{m}\frac{1}{\sigma_i^2}\left(r_i\frac{\sqrt{(x-x_i)^2+(y-y_i)^2}-\frac{(x-x_i)^2}{\sqrt{(x-x_i)^2+(y-y_i)^2}}}{(x-x_i)^2+(y-y_i)^2}-1\right)\\
&= \sum_{i=1}^{m}\frac{1}{\sigma_i^2}\left(\frac{r_i\sin^2\alpha_i}{\sqrt{(x-x_i)^2+(y-y_i)^2}}-1\right)\\
\Rightarrow -E\left[\frac{\partial^2}{\partial x^2}\log f(\mathbf{r};\theta)\right] &= \sum_{i=1}^{m}\frac{1}{\sigma_i^2}\left(1-\frac{E\left[r_i\right]\sin^2\alpha_i}{\sqrt{(x-x_i)^2+(y-y_i)^2}}\right)\\
\Rightarrow I_{xx} &= \sum_{i=1}^{m}\frac{\cos^2\alpha_i}{\sigma_i^2}.
\end{aligned}$$

Similarly, we can show that

$$I_{yy}=\sum_{i=1}^{m}\frac{\sin^2\alpha_i}{\sigma_i^2}.$$

Further,

$$\begin{aligned}
\frac{\partial^2}{\partial y\partial x}\log f(\mathbf{r};\theta) &= \frac{\partial}{\partial y}\sum_{i=1}^{m}\frac{1}{\sigma_i^2}\left(r_i\frac{x-x_i}{\sqrt{(x-x_i)^2+(y-y_i)^2}}-(x-x_i)\right)\\
&= -\sum_{i=1}^{m}\frac{r_i(x-x_i)}{\sigma_i^2}\left(\frac{(y-y_i)}{((x-x_i)^2+(y-y_i)^2)^{\frac{3}{2}}}\right)\\
&= -\sum_{i=1}^{m}\frac{r_i(x-x_i)}{\sigma_i^2}\left(\frac{(y-y_i)}{((x-x_i)^2+(y-y_i)^2)^{\frac{3}{2}}}\right)\\
\Rightarrow -E\left[\frac{\partial^2}{\partial y\partial x}\log f(\mathbf{r};\theta)\right] &= \sum_{i=1}^{m}\frac{E[r_i]}{\sigma_i^2}\left(\frac{(x-x_i)(y-y_i)}{((x-x_i)^2+(y-y_i)^2)^{\frac{3}{2}}}\right)\\
\Rightarrow I_{xy}=I_{yx} &= \sum_{i=1}^{m}\frac{\cos\alpha_i\sin\alpha_i}{\sigma_i^2}.
\end{aligned}$$

To summarize the above results,

$$\begin{aligned}
I_{xx} &= \sum_{i=1}^{m} \frac{1}{\sigma_i^2}\left(\frac{(x-x_i)^2}{(x-x_i)^2+(y-y_i)^2}\right) = \sum_{i=1}^{m} \frac{\cos^2(\alpha_i)}{\sigma_i^2}, \\
I_{xy} = I_{yx} &= \sum_{i=1}^{m} \frac{1}{\sigma_i^2}\left(\frac{(x-x_i)(y-y_i)}{(x-x_i)^2+(y-y_i)^2}\right) = \sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2}, \\
I_{yy} &= \sum_{i=1}^{m} \frac{1}{\sigma_i^2}\left(\frac{(y-y_i)^2}{(x-x_i)^2+(y-y_i)^2}\right) = \sum_{i=1}^{m} \frac{\sin^2(\alpha_i)}{\sigma_i^2}.
\end{aligned}$$

Therefore, the Fisher Information Matrix $\mathbf{I}$ is given by

$$\mathbf{I} = \begin{bmatrix} \sum_{i=1}^{m} \frac{\cos^2(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} \\ \sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^{m} \frac{\sin^2(\alpha_i)}{\sigma_i^2} \end{bmatrix}.$$

The determinant of $\mathbf{I}$ is given by

$$\begin{aligned}
\det \mathbf{I} &= \left(\sum_{i=1}^{m} \frac{\cos^2(\alpha_i)}{\sigma_i^2}\right)\left(\sum_{i=1}^{m} \frac{\sin^2(\alpha_i)}{\sigma_i^2}\right) - \left(\sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2}\right)^2 \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\cos^2(\alpha_i)\sin^2(\alpha_j)}{\sigma_i^2\sigma_j^2} - \sum_{i=1}^{m}\sum_{j=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)\cos(\alpha_j)\sin(\alpha_j)}{\sigma_i^2\sigma_j^2} \\
&= \sum_{i=1}^{m}\sum_{j=1}^{m} \frac{1}{\sigma_i^2\sigma_j^2}\left(\cos^2(\alpha_i)\sin^2(\alpha_j) - \cos(\alpha_i)\sin(\alpha_i)\cos(\alpha_j)\sin(\alpha_j)\right) \\
&= \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2\sigma_j^2}\left(\cos(\alpha_i)\sin(\alpha_j) - \cos(\alpha_j)\sin(\alpha_i)\right) \\
&= \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2\sigma_j^2} \det\left(\begin{bmatrix} \cos(\alpha_i) & \sin(\alpha_i) \\ \cos(\alpha_j) & \sin(\alpha_j) \end{bmatrix}\right).
\end{aligned}$$

This implies that the inverse of $\mathbf{I}$ is given by

$$\mathbf{I}^{-1} = \frac{\begin{bmatrix} \sum_{i=1}^{m} \frac{\sin^2(\alpha_i)}{\sigma_i^2} & -\sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} \\ -\sum_{i=1}^{m} \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^{m} \frac{\cos^2(\alpha_i)}{\sigma_i^2} \end{bmatrix}}{\sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m} \frac{\cos(\alpha_i)\sin(\alpha_j)}{\sigma_i^2\sigma_j^2} \det\left(\begin{bmatrix} \cos(\alpha_i) & \sin(\alpha_i) \\ \cos(\alpha_j) & \sin(\alpha_j) \end{bmatrix}\right)}.$$

The CRLB on the covariance matrix can be expressed as

$$\mathbf{C}_x \geq \mathbf{I}^{-1}, \tag{3.32}$$

which implies that the lower bound on the localization error is

$$\begin{aligned}
\Omega_{\mathbf{x},CRLB}(m) &= \sigma_x^2 + \sigma_y^2 = \mathrm{Tr}\{\mathbf{I}^{-1}\} \\
&= \frac{\sum_{i=1}^m \frac{1}{\sigma_i^2}}{\sum_{i=1}^m \sum_{j=1,j\neq i}^m \frac{\cos(\alpha_i)\sin(\alpha_j)}{\sigma_i^2\sigma_j^2} \det\left(\begin{bmatrix} \cos(\alpha_i) & \sin(\alpha_i) \\ \cos(\alpha_j) & \sin(\alpha_j) \end{bmatrix}\right)} \\
&= \frac{\sum_{i=1}^m \frac{1}{\sigma_i^2}}{\sum_{i=1}^m \sum_{j=1,j\neq i}^m \frac{\cos(\alpha_i)\sin(\alpha_j)}{\sigma_i^2\sigma_j^2} \mathbf{u}_z \cdot (\mathbf{u}_i \times \mathbf{u}_j)} \\
&= \frac{\sum_{i=1}^m \frac{1}{\sigma_i^2}}{\sum_{i=1}^m \sum_{j=1,j>i}^m \left(\frac{\cos(\alpha_i)\sin(\alpha_j)-\cos(\alpha_j)\sin(\alpha_i)}{\sigma_i^2\sigma_j^2}\right) \mathbf{u}_z \cdot (\mathbf{u}_i \times \mathbf{u}_j)} \\
&= \frac{\sum_{i=1}^m \frac{1}{\sigma_i^2}}{\sum_{i=1}^m \sum_{j=1,j>i}^m \frac{\|\mathbf{u}_i \times \mathbf{u}_j\|^2}{\sigma_i^2\sigma_j^2}} = \frac{\sum_{i=1}^m \frac{1}{\sigma_i^2}}{\sum_{i=1}^m \sum_{j=1,j>i}^m \frac{a_{ij}^2}{\sigma_i^2\sigma_j^2}},
\end{aligned}$$

where $\mathbf{u}_i$ represents the unit vector $\frac{\mathbf{x}_i - \mathbf{x}}{\|\mathbf{x}_i - \mathbf{x}\|}$,

$$a_{ij}^2 = \|\mathbf{u}_i \times \mathbf{u}_j\|^2 = \frac{A_{ij}^2}{R_i^2 R_j^2} = \|\mathbf{u}_i\|^2 \|\mathbf{u}_j\|^2 \sin^2(\alpha_{ij}) = \sin^2(\alpha_{ij}) \leq 1.$$

The quantity a_{ij} is essentially the area of the parallelogram formed by the unit vectors $\mathbf{u}_i$ and $\mathbf{u}_j$ as shown in Figure 3.10.

3.9 Appendix 3B: Proof of Theorem 3

Proof of Theorem 3. The generalized GDOP is defined by

$$\begin{aligned}
\Gamma_m &= \frac{\psi_m}{\gamma_m^2} = \frac{\sum_{i=1}^m \sum_{j=1,j>i}^m \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2\sigma_j^2}}{\left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^2} = \frac{\sum_{i=1}^m \sum_{j=1}^m \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2\sigma_j^2}}{2\left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^2} \\
&= \frac{\sum_{i=1}^m \sum_{j=1}^m \left(\frac{1-\cos(2\alpha_i - 2\alpha_j)}{\sigma_i^2\sigma_j^2}\right)}{4\left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^2} = \frac{\left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^2 - \sum_{i=1}^m \sum_{j=1}^m \left(\frac{\cos(2\alpha_i - 2\alpha_j)}{\sigma_i^2\sigma_j^2}\right)}{4\left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^2} \\
&= \frac{1}{4} - \frac{\sum_{i=1}^m \sum_{j=1}^m \frac{\cos(2\alpha_i - 2\alpha_j)}{\sigma_i^2\sigma_j^2}}{4\left(\sum_{i=1}^m \frac{1}{\sigma_i^2}\right)^2}.
\end{aligned}$$

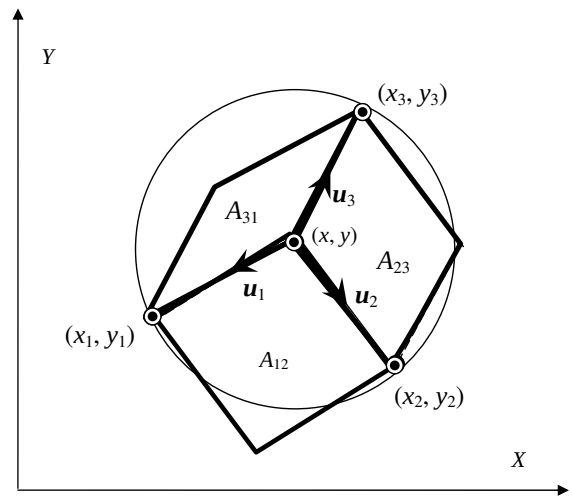


Figure 3.10: Illustration of GDOP and location estimation with $m = 3$ anchor nodes.

When $\alpha_i = \alpha_j$, we see that we obtain the lower bound, $\Gamma_m = 0$. In order to obtain the upper limit, note that

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\cos(2\alpha_i - 2\alpha_j)}{\sigma_i^2\sigma_j^2} &= \sum_{i=1}^{m}\sum_{j=1}^{m}\left(\frac{\cos 2\alpha_i \cos 2\alpha_j + \sin 2\alpha_i \sin 2\alpha_j}{\sigma_i^2\sigma_j^2}\right) \\ &= \left(\sum_{i=1}^{m}\frac{\cos 2\alpha_i}{\sigma_i^2}\right)^2 + \left(\sum_{i=1}^{m}\frac{\sin 2\alpha_i}{\sigma_i^2}\right)^2 \geq 0,\end{aligned}$$

where equality is achieved when

$$\sum_{i=1}^{m}\frac{\cos 2\alpha_i}{\sigma_i^2} = 0,\ \sum_{i=1}^{m}\frac{\sin 2\alpha_i}{\sigma_i^2} = 0. \tag{3.33}$$

Therefore,

$$\Gamma_m = \frac{\psi_m}{\gamma_m^2} = \frac{1}{4} - \frac{\left(\sum_{i=1}^{m}\frac{\cos 2\alpha_i}{\sigma_i^2}\right)^2 + \left(\sum_{i=1}^{m}\frac{\sin 2\alpha_i}{\sigma_i^2}\right)^2}{4\left(\sum_{i=1}^{m}\frac{1}{\sigma_i^2}\right)^2} \leq \frac{1}{4},$$

with equality achieved when (3.33) is satisfied. □

3.10 Appendix 3C: Proof of Theorem 4

Proof of Theorem 4. Without loss of generality, let the mth range estimate be reduced by a factor a, $0 < a \leq 1$. Then

$$\Omega'_{\mathbf{x},CRLB} = \frac{\gamma_{m-1} + \frac{1}{a\sigma_m^2}}{\psi_{m-1} + \frac{1}{a\sigma_m^2}\zeta_{m-1}}, \ \zeta_{m-1} = \sum_{k=1}^{m-1} \frac{\sin^2(\alpha_k - \alpha_m)}{\sigma_k^2}.$$

After some manipulation, it can be shown that

$$\begin{aligned} \zeta_{m-1} &= \frac{\gamma_{m-1}}{2} - \frac{\sqrt{\gamma_{m-1}^2 - 4\psi_{m-1}}\cos(2\alpha_m - 2\nu)}{2} \\ &\geq \frac{\gamma_{m-1}}{2} - \frac{\sqrt{\gamma_{m-1}^2 - 4\psi_{m-1}}}{2}, \end{aligned} \tag{3.34}$$

where the angle ν is defined as

$$\nu = \frac{1}{2}\arctan\left(\frac{\sum_{i=1}^{m-1}\left(\frac{\sin 2\alpha_i}{\sigma_i^2}\right)}{\sum_{i=1}^{m-1}\left(\frac{\cos 2\alpha_i}{\sigma_i^2}\right)}\right). \tag{3.35}$$

The difference between $\Omega_{\mathbf{x},CRLB}$ and $\Omega'_{\mathbf{x},CRLB}$ is then given by

$$\begin{aligned} \Omega_{\mathbf{x},CRLB} - \Omega'_{\mathbf{x},CRLB} &= \frac{\gamma_{m-1} + \frac{1}{\sigma_m^2}}{\psi_{m-1} + \frac{1}{\sigma_m^2}\zeta_{m-1}} - \frac{\gamma_{m-1} + \frac{1}{a\sigma_m^2}}{\psi_{m-1} + \frac{1}{a\sigma_m^2}\zeta_{m-1}} \\ &= \frac{\sigma_m^2\gamma_{m-1} + 1}{\sigma_m^2\psi_{m-1} + \zeta_{m-1}} - \frac{a\sigma_m^2\gamma_{m-1} + 1}{a\sigma_m^2\psi_{m-1} + \zeta_{m-1}} \\ &= \frac{(a-1)\sigma_m^2\psi_{m-1} + (1-a)\sigma_m^2\gamma_{m-1}\zeta_{m-1}}{(\sigma_m^2\psi_{m-1} + \zeta_{m-1})(a\sigma_m^2\psi_{m-1} + \zeta_{m-1})} \\ &= \frac{(1-a)\sigma_m^2(\gamma_{m-1}\zeta_{m-1} - \psi_{m-1})}{(\sigma_m^2\psi_{m-1} + \zeta_{m-1})(a\sigma_m^2\psi_{m-1} + \zeta_{m-1})} \\ &\geq \frac{(1-a)\sigma_m^2\left(\gamma_{m-1}\left(\frac{\gamma_{m-1}}{2} - \frac{\sqrt{\gamma_{m-1}^2 - 4\psi_{m-1}}}{2}\right) - \psi_{m-1}\right)}{(\sigma_m^2\psi_{m-1} + \zeta_{m-1})(a\sigma_m^2\psi_{m-1} + \zeta_{m-1})} \\ &= \frac{\frac{(1-a)\sigma_m^2\gamma_{m-1}^2}{4}\left(1 - \sqrt{1 - \frac{4\psi_{m-1}}{\gamma_{m-1}^2}}\right)^2}{(\sigma_m^2\psi_{m-1} + \zeta_{m-1})(a\sigma_m^2\psi_{m-1} + \zeta_{m-1})} \geq 0, \end{aligned}$$

since $\psi_{m-1} \leq \frac{\gamma_{m-1}^2}{4}$ and $0 < a \leq 1$, and all other quantities are positive. Therefore $\Omega_{\mathbf{x},CRLB} \geq \Omega'_{\mathbf{x},CRLB}$. $\square$

3.11 Appendix 3D: Proof of Theorem 5

Proof of Theorem 5. Given range estimates from m anchor nodes, we know that the location error is given by the relation:

$$\Omega_{\mathbf{x},CRLB}(m) = \frac{\sum_{i=1}^{m} \frac{1}{\sigma_i^2}}{\sum_{i=1}^{m} \sum_{j=1,j>i}^{m} \frac{\sin^2(\alpha_i-\alpha_j)}{\sigma_i^2\sigma_j^2}} = \frac{\gamma_m}{\psi_m}.$$

We would like to analyze the effect of adding a new node $(\alpha_{m+1}, \sigma_{m+1}^2)$:

$$\begin{aligned}
\Omega_{\mathbf{x},CRLB}(m+1) &= \frac{\sum_{i=1}^{m+1} \frac{1}{\sigma_i^2}}{\sum_{i=1}^{m+1} \sum_{j=1,j>i}^{m+1} \frac{\sin^2(\alpha_i-\alpha_j)}{\sigma_i^2\sigma_j^2}} \\
&= \frac{\gamma_m + \frac{1}{\sigma_{m+1}^2}}{\psi_m + \frac{1}{\sigma_{m+1}^2} \sum_{i=1}^{m} \frac{\sin^2(\alpha_i-\alpha_{m+1})}{\sigma_i^2}} = \frac{\sigma_{m+1}^2\gamma_m + 1}{\sigma_{m+1}^2\psi_m + \zeta}, \qquad (3.36)
\end{aligned}$$

where we define

$$\zeta = \sum_{i=1}^{m} \frac{\sin^2(\alpha_i - \alpha_{m+1})}{\sigma_i^2}.$$

Expanding the RHS, we have

$$\begin{aligned}
\zeta &= \sum_{i=1}^{m} \frac{\sin^2(\alpha_i - \alpha_{m+1})}{\sigma_i^2} \qquad (\geq 0) \\
&= \sum_{i=1}^{m} \frac{1-\cos(2\alpha_i - 2\alpha_{m+1})}{2\sigma_i^2} = \frac{\gamma_m}{2} - \frac{1}{2}\sum_{i=1}^{m} \frac{\cos(2\alpha_i - 2\alpha_{m+1})}{\sigma_i^2} \\
&= \frac{\gamma_m}{2} - \frac{1}{2}\sum_{i=1}^{m} \left(\frac{\cos 2\alpha_i \cos 2\alpha_{m+1} + \sin 2\alpha_i \sin 2\alpha_{m+1}}{\sigma_i^2} \right) \\
&= \frac{\gamma_m}{2} - \frac{\cos 2\alpha_{m+1}}{2} \sum_{i=1}^{m} \left(\frac{\cos 2\alpha_i}{\sigma_i^2} \right) - \frac{\sin 2\alpha_{m+1}}{2} \sum_{i=1}^{m} \left(\frac{\sin 2\alpha_i}{\sigma_i^2} \right) \\
&= \frac{\gamma_m}{2} - \frac{\eta \cos 2\alpha_{m+1}}{2} \underbrace{\sum_{i=1}^{m} \frac{\left(\frac{\cos 2\alpha_i}{\sigma_i^2}\right)}{\eta}}_{\cos(2\nu)} - \frac{\sin 2\alpha_{m+1}}{2} \underbrace{\sum_{i=1}^{m} \frac{\left(\frac{\sin 2\alpha_i}{\sigma_i^2}\right)}{\eta}}_{\sin(2\nu)},
\end{aligned}$$

where

$$\begin{aligned}\eta &= \sqrt{\left(\sum_{i=1}^{m}\frac{\cos 2\alpha_i}{\sigma_i^2}\right)^2+\left(\sum_{i=1}^{m}\frac{\sin 2\alpha_i}{\sigma_i^2}\right)^2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}\frac{\cos\left(2\alpha_i-2\alpha_j\right)}{\sigma_i^2\sigma_j^2}}\\ &= \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{m}\left(\frac{1-2\sin^2\left(\alpha_i-\alpha_j\right)}{\sigma_i^2\sigma_j^2}\right)} = \sqrt{\gamma_m^2-4\psi_m}.\end{aligned}$$

Defining the angle ν as

$$\nu = \frac{1}{2}\arctan\left(\frac{\sum_{i=1}^{m}\frac{\sin 2\alpha_i}{\sigma_i^2}}{\sum_{i=1}^{m}\frac{\cos 2\alpha_i}{\sigma_i^2}}\right),$$

we have

$$\zeta = \frac{\gamma_m}{2}-\frac{\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}. \tag{3.37}$$

From (3.36) and (3.37),

$$\Omega_{\mathbf{x},CRLB}(m+1) = \frac{\sigma_{m+1}^2\gamma_m+1}{\sigma_{m+1}^2\psi_m+\zeta} = \frac{\sigma_{m+1}^2\gamma_m+1}{\sigma_{m+1}^2\psi_m+\frac{\gamma_m}{2}-\frac{\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}}.$$

We see that $\Omega_{\mathbf{x},CRLB}(m+1)$ is minimized when ζ is maximized. From (3.37), this implies that the optimal value of α_{m+1} is $\alpha_{m+1}^{\star} = \nu \pm \frac{\pi}{2}$ (Corollary 5).

In order to compute the improvement due to the addition of a new node, we look at

$$\begin{aligned}\Omega_{\mathbf{x},CRLB}(m)-\Omega_{\mathbf{x},CRLB}(m+1) &= \frac{\gamma_m}{\psi_m}-\frac{\sigma_{m+1}^2\gamma_m+1}{\sigma_{m+1}^2\psi_m+\frac{\gamma_m}{2}-\frac{\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}}\\ &= \frac{\frac{\gamma_m^2}{2}-\frac{\gamma_m\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}-\psi_m}{\psi_m\left(\sigma_{m+1}^2\psi_m+\frac{\gamma_m}{2}-\frac{\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}\right)}\\ &= \gamma_m^2\left[\frac{\frac{1}{2}-\frac{\sqrt{1-4\frac{\psi_m}{\gamma_m^2}}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}-\frac{\psi_m}{\gamma_m^2}}{\psi_m\left(\sigma_{m+1}^2\psi_m+\frac{\gamma_m}{2}-\frac{\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}\right)}\right]\\ &= \left[\frac{\frac{1}{2}-\frac{\sqrt{1-4\Gamma_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}-\Gamma_m}{\Gamma_m\left(\sigma_{m+1}^2\psi_m+\frac{\gamma_m}{2}-\frac{\sqrt{\gamma_m^2-4\psi_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2}\right)}\right].\end{aligned}$$

Since the denominator is always positive, to verify whether there is an improvement in the location error or not, we verify that:

$$\begin{aligned} & \frac{1}{2} - \frac{\sqrt{1-4\Gamma_m}\cos\left(2\alpha_{m+1}-2\nu\right)}{2} - \Gamma_m \\ = & \frac{\left(\sqrt{1-4\Gamma_m} - \cos\left(2\alpha_{m+1}-2\nu\right)\right)^2 + \sin^2\left(2\alpha_{m+1}-2\nu\right)}{2} \\ \geq & \ 0, \end{aligned}$$

where equality holds when $\Gamma_m = 0$ and $\alpha_{m+1} = \nu$. Therefore, $\Omega_{\mathbf{x},CRLB}(m) \geq \Omega_{\mathbf{x},CRLB}(m+1)$. □

3.12 Appendix 3E: CRLB for exponentially distributed range estimates

The joint probability density function of m i.i.d. exponentially distributed range estimates with means λ_i, $i = 1, 2, \cdots, m$ is given by

$$\begin{aligned} f(\mathbf{r};\theta) &= \prod_{i=1}^{m} \frac{1}{\lambda_i} \exp\left(-\frac{(r_i - \|\mathbf{x}-\mathbf{x}_i\|)}{\lambda_i}\right) \\ \Rightarrow \log f(\mathbf{r};\theta) &= -\sum_{i=1}^{m} \log\lambda_i - \sum_{i=1}^{m} \frac{(r_i - \|\mathbf{x}-\mathbf{x}_i\|)}{\lambda_i}. \end{aligned}$$

The Fisher information matrix is given by

$$\mathbf{I} = \begin{bmatrix} I_{xx} & I_{xy} \\ I_{yx} & I_{yy} \end{bmatrix}$$

where

$$\begin{aligned} I_{xx} &= -E\left[\frac{\partial^2}{\partial x^2}\log f(\mathbf{r};\theta)\right], \ I_{xy} = -E\left[\frac{\partial^2}{\partial x \partial y}\log f(\mathbf{r};\theta)\right] \\ I_{yx} &= -E\left[\frac{\partial^2}{\partial y \partial x}\log f(\mathbf{r};\theta)\right], \ I_{yy} = -E\left[\frac{\partial^2}{\partial y^2}\log f(\mathbf{r};\theta)\right] \end{aligned}$$

We obtain

$$\begin{aligned} \frac{\partial}{\partial x}\log f(\mathbf{r};\theta) &= -\sum_{i=1}^{m} \frac{1}{\lambda_i}\frac{\partial}{\partial x}\left(r_i - \sqrt{(x-x_i)^2 + (y-y_i)^2}\right) \\ &= \sum_{i=1}^{m} \frac{1}{\lambda_i}\frac{(x-x_i)}{\sqrt{(x-x_i)^2 + (y-y_i)^2}} \end{aligned}$$

From the above equation, we obtain

$$\begin{aligned}
\frac{\partial^2}{\partial x^2}\log f(\mathbf{r};\theta) &= \frac{\partial}{\partial x}\sum_{i=1}^{m}\frac{1}{\lambda_i}\frac{(x-x_i)}{\sqrt{(x-x_i)^2+(y-y_i)^2}} \\
&= \sum_{i=1}^{m}\frac{1}{\lambda_i}\left(\frac{\sqrt{(x-x_i)^2+(y-y_i)^2}-\frac{(x-x_i)^2}{\sqrt{(x-x_i)^2+(y-y_i)^2}}}{(x-x_i)^2+(y-y_i)^2}\right) \\
&= \sum_{i=1}^{m}\frac{1}{\lambda_i}\left(\frac{\sin^2\alpha_i}{R_i}\right) \\
\Rightarrow I_{xx} &= -\sum_{i=1}^{m}\frac{\sin^2\alpha_i}{R_i\lambda_i}.
\end{aligned}$$

Similarly, we can show that

$$I_{yy} = -\sum_{i=1}^{m}\frac{\cos^2\alpha_i}{R_i\lambda_i}.$$

Further,

$$\begin{aligned}
\frac{\partial^2}{\partial y\partial x}\log f(\mathbf{r};\theta) &= \frac{\partial}{\partial y}\sum_{i=1}^{m}\frac{1}{\lambda_i}\frac{(x-x_i)}{\sqrt{(x-x_i)^2+(y-y_i)^2}} \\
&= -\sum_{i=1}^{m}\frac{(x-x_i)}{\lambda_i}\left(\frac{(y-y_i)}{((x-x_i)^2+(y-y_i)^2)^{\frac{3}{2}}}\right) \\
\Rightarrow I_{xy}=I_{yx} &= \sum_{i=1}^{m}\frac{\cos\alpha_i\sin\alpha_i}{\lambda_i R_i}.
\end{aligned}$$

Therefore, the Fisher Information Matrix $\mathbf{I}$ is given by

$$\mathbf{I} = \begin{bmatrix} -\sum_{i=1}^{m}\frac{\sin^2\alpha_i}{R_i\lambda_i} & \sum_{i=1}^{m}\frac{\cos\alpha_i\sin\alpha_i}{\lambda_i R_i} \\ \sum_{i=1}^{m}\frac{\cos\alpha_i\sin\alpha_i}{\lambda_i R_i} & -\sum_{i=1}^{m}\frac{\cos^2\alpha_i}{R_i\lambda_i} \end{bmatrix}.$$

Proceeding in a manner similar to Appendix 3A, it is straightforward to show that the localization error expressions for the unbiased Gaussian and the exponential cases, if $\sigma_i^2 = \lambda_i R_i$.

Chapter 4

The NLOS problem in UWB PoLoNets

4.1 Introduction

As discussed in Chapter 2, in PoLoNets deployed in dense multipath propagation environments, especially indoors or in urban scenarios, the LOS path between nodes may be obstructed. As a result, in NLOS conditions, TOA-based range estimates are positively biased with high probability, since the first arriving multipath component travels a distance that is in excess of the true LOS distance, as illustrated in Figure 2.4. A similar effect is seen in the case of RSS-based range estimates, where the received signal power is reduced due to the obstruction of the LOS path. These effects result in range estimates that are often much larger than the true distances and therefore, in NLOS scenarios, the accuracy of node location estimates can be adversely affected.

The problem of location-estimation with biased NLOS range estimates has been considered before, mostly in the context of cellular communications [24], [4], where it has been shown that the bias errors in NLOS range estimates lead to large errors in the computation of a node's location. Similar observations have been made with UWB signals [43], where it was demonstrated that the NLOS bias errors can be on the order of several meters and are much larger than the range measurement errors in LOS scenarios.

Broadly speaking, the literature on the NLOS problem falls in two categories: NLOS identification and NLOS mitigation. The former deals with the problem of distinguishing between LOS and NLOS range estimates, whereas the latter typically deals with the reduction of the adverse impact of NLOS range errors on the accuracy of location-estimates, assuming the NLOS range estimates have been identified.

4.1.1 NLOS Identification

Statistical NLOS identification techniques for cellular systems have been discussed previously [49], [50] that rely on a time-series of range measurements. A comprehensive discussion of a decision-theoretic framework for NLOS identification based on the statistical distributions of range estimates in LOS and NLOS scenarios was presented in [50].

While such approaches [49], [50], can be directly extended to UWB PoLoNets, (i) for PoLoNets with low or no mobility, using a time-series of range estimates may not be feasible, and (ii) specific features of UWB multipath channels can be exploited to enhance the accuracy of NLOS identification. In particular, UWB signals possess higher temporal resolution and robustness to multipath fading than narrowband and wideband signals. These characteristics create considerable differences in the behavior of (a) the Time-of-Arrival (TOA), (b) Received signal-strength (RSS) and (c) the temporal energy dispersion, of the received signal in different propagation scenarios. The temporal dispersion of received signal energy in indoor UWB propagation channels has been extensively characterized (see, for instance, [1, 51]), and is quantified through the delay-spread statistics, such as the root-mean-squared delay spread (RDS) [1]. The variations of the TOA, RSS and RDS estimates from UWB signals can be utilized to distinguish between LOS and NLOS propagation scenarios.

In this chapter, we formulate a NLOS identification technique based on the joint statistical hypothesis testing of a *single set* of TOA, RSS and RDS estimates from a given received UWB signal. We initially assume that the statistical models for the TOA, RSS and RDS estimates are known *a priori* from a statistical characterization of the propagation environment. Since the distributions of these parameters are known to vary with the distance [1], [51], the distance is modeled as a nuisance parameter that needs to be estimated. Given a distance estimate, we show that combinations of conditional probabilities of the estimated values of the TOA, RSS and RDS can provide reliable estimates of the state of the channel (LOS or NLOS).

We show that for the channel parameters extracted from the UWB measurements discussed in [1], the use of RDS estimates provides the most accurate means of identifying the channel state, with a probability of error on the order of 1%. Further, we show that delay spread statistics of the received signal remain a reliable means of channel state identification when (a) the complete statistical characterization of the propagation environment is unknown *a priori*, and (b) when an accurate estimate of the distance is not available. Performance metrics that allow the evaluation of the proposed techniques in other propagation scenarios with different statistical models are provided. The performance of the NLOS identification based on delay spread statistics, from the perspective of localization accuracy, is investigated in a location-tracking experiment with indoor UWB measurements in Chapter 6.

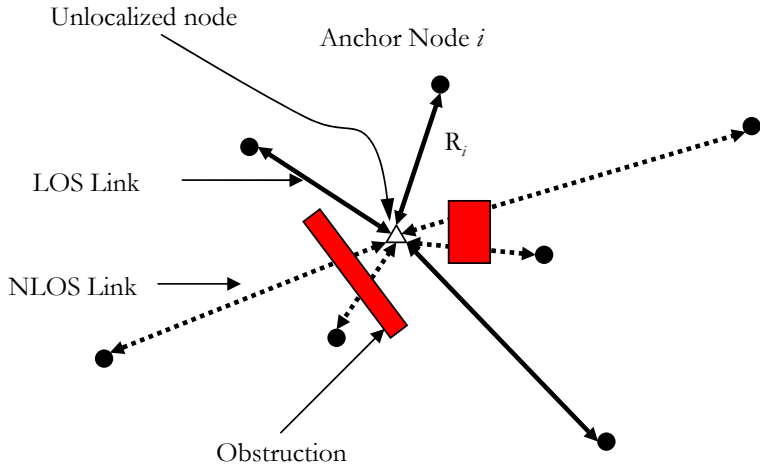


Figure 4.1: The NLOS Problem: In the absence of LOS in TOA-based ranging between two nodes A and B, the estimated range is larger than the true distance between A and B.

4.1.2 NLOS Mitigation

Once we are able to accurately distinguish between LOS and NLOS range estimates, we can then turn our attention to the problem of NLOS mitigation. Equipped with the ability to distinguish between LOS and NLOS range estimates, the Cramer-Rao Lower Bound (CRLB) analysis presented in [52], [53], characterized the performance of the minimum variance unbiased estimator (MVUE) [34] of a node's location, given a mixture of (unbiased) LOS and (biased) NLOS range estimates. This analysis showed that, in the absence of *a priori* statistical information on the NLOS range estimates, the MVUE discards the biased NLOS range estimates and utilizes only LOS range information while estimating sensor locations. However, as will be demonstrated in the following sections, in the case of practical non-efficient [34] estimators such as the commonly used Least-Squares (LS) estimator [20], discarding NLOS range information does not necessarily improve performance. Additionally, for two (three) dimensional location-estimation, the LS estimator requires at least three (four) range estimates in order to obtain an unambiguous solution. In ad hoc networks, limited connectivity with anchors may imply that we may not have the luxury of discarding any range estimates. This suggests that in general, given a mixture of LOS and NLOS range estimates, we may be required to use the entire set of available range information in order to compute a node's location, as shown in Figure 4.1.

A Semi-Definite Programming (SDP) approach to node localization based on connectivity information was investigated in [54], and a quadratic programming approach with NLOS range estimates was discussed in [55], but these approaches result in high computational complexity [56]. The Residual Weighting Algorithm (Rwgh) was proposed in [57], whose advantage is that NLOS identification is not required *a priori*. However, this algorithm

implicitly assumes that the range measurement noise is much smaller than the NLOS bias introduced, in order to inherently distinguish between the residuals from LOS and NLOS range estimates. More importantly, it relies on the availability of a large number of range estimates, several of which are LOS, so that the set of range estimates finally selected to compute a node's location results in the smallest residual error. However, in indoor networks, situations may arise where *only NLOS* range estimates are available while estimating a node's location.

The statistics of the NLOS bias errors depend on the spatial distribution of scatterers in the propagation environment. Given the scattering model for the environment, the statistics of TOA measurements can be obtained and well-known techniques such as maximum *a posteriori* (MAP) and maximum-likelihood (ML) estimation can be employed to mitigate the effects of NLOS errors [58]. However, in an ad hoc wireless network, the distribution of scatterers in the propagation environment may not be available and therefore, we assume that the statistics of the NLOS bias errors are unknown *a priori*.

In this chapter, we present a novel computationally-efficient linear programming (LP) approach that effectively incorporates both LOS and NLOS range information into the estimation of a node's location. A linear programming approach was briefly mentioned for the case of NLOS range estimates in [56], but was not pursued. We demonstrate that this low-complexity LP approach can be generalized to handle a mixture of LOS and NLOS range estimates (with the "only-LOS" and "only-NLOS" range information scenarios as sub-cases) without discarding any range information. The main advantages of this approach are:

- The statistics of the NLOS bias errors are not assumed to be known [15] *a priori*,
- No range information is discarded,
- This method outperforms the LS estimator, given a mixture of LOS and NLOS range estimates,
- This method can be generalized and extended to handle degenerate cases with insufficient (< 3) LOS range estimates, and
- Can be extended to include other forms of constraints.

In the proposed approach, we leverage the following features of UWB TOA-based range estimation: (a) the range bias errors in NLOS conditions are always positive and significantly larger in magnitude than the range measurement errors in LOS conditions [43], and (b) NLOS range estimates are readily distinguished from LOS range estimates through channel identification [26]. It is important to point out that although the development of the approach is from the perspective of UWB PoLoNets, the approach itself is generally valid for any location-aware system with the mentioned features. For the purpose of analytical simplicity and clear exposition, the development and insights that follow largely pertain

to two-dimensional location estimation, but can easily be extended to three-dimensional localization scenarios.

This chapter is organized in two parts. Part I focuses on the NLOS Identification using received signal statistics. The organization of Part I is as follows: in Section 4.2, we discuss the modeling of the statistics of the TOA, RSS and RDS estimates in indoor LOS and NLOS environments based on the UWB measurement campaign discussed in [1]. The formulation of the statistical hypothesis-testing method is outlined in Section 4.3, and the corresponding simulation results are presented in Section 4.4. Part I concludes in Section 4.5.

Part II discusses the use of the linear programming for NLOS mitigation. In section 4.6, we discuss the impact of NLOS bias errors on the accuracy of LS location estimates and the motivation for new approaches. Section 4.7 discusses the LP approach to incorporating LOS range estimates, NLOS range estimates and a mixture of LOS and NLOS range estimates into node location estimation. In section 4.8, we discuss some extensions and a series of sub-cases that need to be addressed in order to generalize the proposed approach. Simulation results are presented in Section 4.9, where we evaluate the performance of the proposed method in terms of node localization accuracy. We also present two practical scenarios (including a 3D indoor scenario) in Section 4.10 where the efficacy of the proposed approach is highlighted. Our conclusions are presented in Section 4.11.

Original Contributions included in this chapter are:

- The complete development and analysis of a novel NLOS identification method based on received signal statistics [26]: We statistically characterize the ability of TOA, RSS and RMS delay spread estimates to distinguish between LOS and NLOS propagation based on an extensive indoor measurement campaign. Simulation and measurement results are used to demonstrate that the delay spread statistics of UWB signals can, even in isolation and without complete statistical information, serve as robust and computationally-efficient indicators of the LOS/NLOS nature of propagation.

- The development and analysis of a novel NLOS mitigation technique based on linear programming [27], [28], that (i) allows us to incorporate NLOS range information into location-estimation, but (ii) does not allow NLOS bias errors to degrade node localization accuracy.

Part I: NLOS Identification based on Received Signal Statistics

4.2 Statistics of the UWB TOA, RSS and RDS Estimates in Indoor Environments

The formulation of the NLOS identification problem is as follows: we would like to identify the state H of the channel between the transmit and receive nodes, separated physically by a distance d, given estimates of the TOA, RSS and RDS observed at the receive node. Here, the channel state $H = H_0$ corresponds to LOS propagation and $H = H_1$ corresponds to NLOS propagation. NLOS propagation has popularly been classified as "soft" ("obstructed") NLOS, where the LOS multipath component is present albeit attenuated, and "hard" NLOS, where the LOS path is severely attenuated or absent. From a localization standpoint, the soft-NLOS cases are classified as LOS scenarios, since the TOA of the LOS multipath component can still be estimated for ranging purposes, i.e., range estimates are not necessarily biased in soft-NLOS propagation environments. In the following development, we present statistical models for the TOA, RSS and RDS estimates in an indoor UWB channel, considering both LOS and NLOS scenarios.

TOA Estimates

The received UWB signal, neglecting non-linear effects, can be modeled as a function of time τ:

$$r(\tau) = \sum_{k=0}^{L-1} \alpha_k p(\tau - \tau_k) + n(\tau),$$

where $p(\tau)$ is the received (LOS) pulse shape, α_k, $k = 0, 1, \cdots, L-1$ denote the L multipath coefficients, and $n(\tau)$ represents additive white Gaussian noise at the receiver. Here, $\tau_0 = \frac{d}{c}$ represents the true TOA of the signal, where c denotes the speed of light. In LOS propagation conditions, the estimate of the TOA $\hat{\tau}_0$ is modeled as [18], [43] an unbiased Gaussian random variable:

$$\hat{\tau}_0 = \frac{d}{c} + n_\tau, \qquad (H = H_0)$$

where n_τ is the zero-mean Gaussian TOA measurement noise. The corresponding density function of $\hat{\tau}_0$, conditioned on the distance d is then given by:

$$f_{\hat{\tau}_0}(\tau \,|d,\ H_0) = \left(2\pi\sigma_\tau^2\right)^{-1/2} \exp\left(-\frac{(\tau - \frac{d}{c})^2}{2\sigma_\tau^2}\right),$$

where σ_τ represents the standard deviation of the TOA measurement noise. Since the minimum variance of an unbiased TOA estimate is known to be inversely proportional to the signal-to-noise ratio [17], the standard deviation of the TOA estimates σ_τ is dependent on the distance d, and is modeled as:

$$\sigma_\tau = K_\tau d^{\frac{\beta}{2}},$$

where K_τ is a proportionality constant that depends on physical layer parameters, and β is the path-loss exponent.

In a NLOS propagation scenario, the multipath component corresponding to the true distance d between the nodes is either absent or severely attenuated, and the TOA estimate $\hat{\tau}_0$ is positively biased with high probability. The model for $\hat{\tau}_0$ in NLOS conditions is therefore the superposition of an unbiased Gaussian measurement noise and a random positive bias:

$$\hat{\tau}_0 = \frac{d}{c} + n_\tau + b_\tau, \qquad (H = H_1)$$

where b_τ is the NLOS bias. The NLOS bias b_τ is assumed to be independent of the range measurement noise and has been modeled as an exponential random variable [43] with mean λ_b. Based on these assumptions, as shown in Appendix 4A, the density function of $\hat{\tau}_0$, conditioned on the distance d, can be written as:

$$f_{\hat{\tau}_0}(\tau\,|d,\ H_1) = \frac{1}{2\lambda_b}\exp\left(\frac{\sigma_\tau^2}{2\lambda_b^2}\right)\exp\left(-\frac{(\tau-\frac{d}{c})}{\lambda_b}\right)\left[1+\operatorname{erf}\left(\frac{\lambda_b(\tau-\frac{d}{c})-\sigma_\tau^2}{\sqrt{2}\sigma_\tau\lambda_b}\right)\right].$$

Summarizing the above results, the density functions of the TOA estimate $\hat{\tau}_0$ for the two propagation scenarios can be written as

$$f_{\hat{\tau}_0}(\tau\,|d,\ H) = \begin{cases} (2\pi\sigma_\tau^2)^{-1/2}\exp\left(-\frac{(\tau-\frac{d}{c})^2}{2\sigma_\tau^2}\right), & H = H_0, \\ \frac{1}{2\lambda_b}\exp\left(\frac{\sigma_\tau^2}{2\lambda_b^2}\right)\exp\left(-\frac{(\tau-\frac{d}{c})}{\lambda_b}\right)\left[1+\operatorname{erf}\left(\frac{\lambda_b(\tau-\frac{d}{c})-\sigma_\tau^2}{\sqrt{2}\sigma_\tau\lambda_b}\right)\right], & H = H_1. \end{cases} \tag{4.1}$$

In the above equations,

$$\sigma_\tau = \begin{cases} K_\tau d^{\frac{\beta_L}{2}}, & H = H_0, \\ K_\tau d^{\frac{\beta_N}{2}}, & H = H_1, \end{cases}$$

where β_L and β_N represent the path-loss exponents in LOS and NLOS scenarios respectively.

RSS Estimates

The normalized RSS (in dB) S_{dB} is defined as:

$$S_{dB} = 10\log_{10}\left(\frac{1}{P_0 T}\int_T |r(\tau)|^2 d\tau\right),$$

where P_0 is the received power measured at a reference distance d_0, and T is the measurement interval for the received signal. Due to the absence of significant multipath fading, the estimated RSS has been modeled as a lognormal random variable [1], [51], [59]:

$$\hat{S}_{dB} = 10\beta \log_{10}\left(\frac{d}{d_0}\right) + X,$$

where β is the path loss exponent, and X is a zero-mean Gaussian random variable with different variances in LOS and NLOS scenarios:

$$f_X(x) = \begin{cases} \left(2\pi\sigma_{SL}^2\right)^{-1/2} \exp\left(-\frac{x^2}{2\sigma_{SL}^2}\right), & H = H_0, \\ \left(2\pi\sigma_{SN}^2\right)^{-1/2} \exp\left(-\frac{x^2}{2\sigma_{SN}^2}\right), & H = H_1. \end{cases}$$

Therefore, in LOS and NLOS scenarios, the conditional density functions of $\hat{S}_{dB}$ can be succinctly represented by:

$$f_{\hat{S}_{dB}}(s\,|d,\ H\,) = \begin{cases} \left(2\pi\sigma_{SL}^2\right)^{-1/2} \exp\left(-\frac{(s-10\beta_L \log_{10}(d))^2}{2\sigma_{SL}^2}\right), & H = H_0, \\ \left(2\pi\sigma_{SN}^2\right)^{-1/2} \exp\left(-\frac{(s-10\beta_N \log_{10}(d))^2}{2\sigma_{SN}^2}\right), & H = H_1. \end{cases} \tag{4.2}$$

RDS Estimates

Given the received signal $r(\tau)$, and the received LOS pulse $p(\tau)$, the temporal channel impulse response $h_c(\tau)$ can be estimated using a deconvolution algorithm such as the CLEAN algorithm [60], which has previously been applied to impulse-based UWB measurements [60], [61]. The estimated channel impulse response $\hat{h}_c(\tau)$ is a superposition of $\hat{L}$ discrete multipath components:

$$\hat{h}_c(\tau) = \sum_{k=0}^{\hat{L}-1} \hat{\alpha}_k \delta(\tau - \hat{\tau}_k), \tag{4.3}$$

where $\hat{\alpha}_k$ and $\hat{\tau}_k$ are the estimated amplitude (including polarity) and arrival-time of the kth multipath component respectively. The estimated number of multipath components $\hat{L}$ is determined by the relative strengths of the multipath components and the amplitude threshold applied by the deconvolution algorithm. The deconvolution algorithm typically retains those multipath components whose amplitudes are within a certain threshold of the strongest multipath component. The most common metric used to quantify the dispersion of the received signal energy over time by the channel is the RDS $\hat{\tau}_{rms}$, which is defined as:

$$\hat{\tau}_{rms} = \sqrt{\frac{\sum_k \hat{\alpha}^2 \hat{\tau}_k^2}{\sum_k \hat{\alpha}_k^2} - \hat{\tau}_m^2}, \tag{4.4}$$

where $\hat{\tau}_m$ is the mean excess delay of the channel, defined as $\hat{\tau}_m = \frac{\sum_k \hat{\alpha}_k^2 \tau_k}{\sum_k \hat{\alpha}_k^2}$. Figure 4.2 illustrates the dispersion in the received multipath energy over time in two sample LOS and

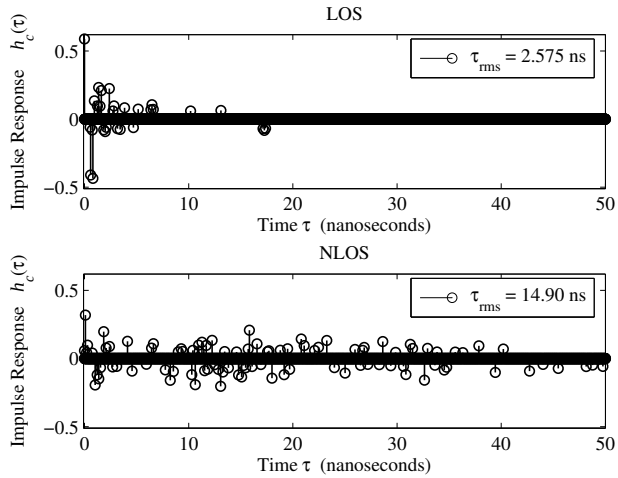


Figure 4.2: Sample LOS (top) and NLOS (bottom) channel impulse responses in an indoor office propagation measurement [1]. Both channel impulse responses were computed using the CLEAN algorithm with a threshold of 20 dB, and were subsequently normalized to unit energy. The RMS delay spread and the number of significant multipath components in the LOS and NLOS cases are respectively 2.575 ns and 25 (top), and 14.90 ns and 76 (bottom).

NLOS channel impulse responses extracted from measurements [1]. In the LOS case, the energy is seen to be concentrated in the first-arriving path, whereas in the NLOS case, the energy is well distributed over several multipath components. We see that this difference is captured by the values of $\hat{\tau}_{rms}$ computed for the two cases. Further, we also observe that for a fixed amplitude threshold (20 dB in this case), the estimated number of multipath components $\hat{L}$ is much larger in the NLOS case than the LOS case.

The RDS estimate has been observed to be positive and Gaussian distributed in several measurement campaigns [1, 51, 59]. Further, it has been observed [1, 51, 59, 62] that the average RDS increases with the distance between the nodes. The rationale behind this trend is that the average number of significant scatterers in a typical indoor propagation environment increases as the distance between the transmit and receive antennas increases. The mean of the RDS estimate (in nanoseconds) is assumed to vary linearly with distance [1, 51, 62] as:

$$\mu_{\hat{\tau}_{rms}}(d) = \begin{cases} a_L(d-d_0)+b_L, & H=H_0, \\ a_N(d-d_0)+b_N, & H=H_1, \end{cases}$$

and we assume that the variance of the RDS estimates $\sigma^2_{\hat{\tau}_{rms}}$ is independent of distance. With these assumptions, the distribution of the RDS estimate can be written as

$$f_{\hat{\tau}_{rms}}(\tau; d\,|H\,) = \begin{cases} c_L\left(2\pi\sigma^2_{\hat{\tau}_{rms}L}\right)^{-1/2}\exp\left(-\frac{(\tau-(a_L(d-d_0)+b_L))^2}{2\sigma^2_{\hat{\tau}_{rms}L}}\right)u(\tau), & H=H_0, \\ c_N\left(2\pi\sigma^2_{\hat{\tau}_{rms}N}\right)^{-1/2}\exp\left(-\frac{(\tau-(a_N(d-d_0)+b_N))^2}{2\sigma^2_{\hat{\tau}_{rms}N}}\right)u(\tau), & H=H_1, \end{cases} \tag{4.5}$$

Table 4.1:]
Parameters obtained from UWB Measurements [1]

Measurement Parameters	LOS $H = H_0$	NLOS $H = H_1$
β	1.8	2.5
λ_b	-	4.5 ns
σ_S	2.75	4.1
a	1.42	15.30
b	0.03	0.15
$\sigma_{\tau_{rms}}$	2.17	2.41

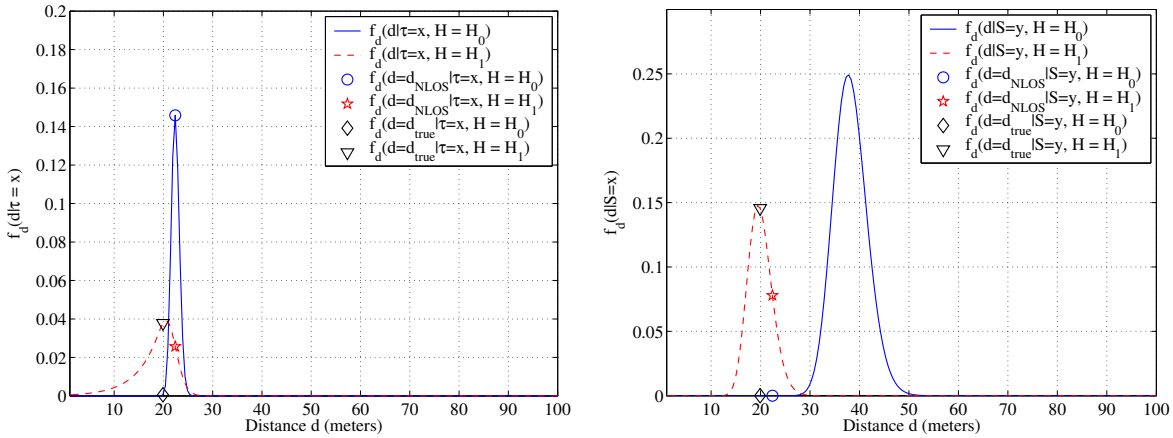
where $u(\cdot)$ is the Heaviside function, and $\{c_L, c_N\}$ are normalization constants that ensure that the area under the above probability density functions is unity (see Appendix 4B).

The parameters corresponding to above the models for the TOA, RSS and RDS estimates extracted from the measurements in [1] are presented in Table 4.1. In the following section, we discuss a statistical NLOS identification method based on the conditional distributions of the TOA, RSS and RDS estimates given in (4.1), (4.2) and (4.5) respectively.

4.3 Statistical NLOS Identification

Suppose we are given a set of the TOA, the RSS and the RDS estimates: $\mathcal{S} = \{x_1, x_2, x_3\}$, where x_1, x_2 and x_3 are realizations of $\hat{\tau}_0$, $\hat{S}_{dB}$ and $\hat{\tau}_{rms}$ respectively, obtained from the received signal. As discussed in the previous section, the conditional distributions of the TOA, RSS and RDS estimates are functions of the distance d and the channel state H, given by (4.1), (4.2) and (4.5) respectively. Provided the physical distance d between the transmit and receive nodes is known exactly, the state of the channel can be identified by comparing the likelihood values for each of the estimates x_i, conditioned on the distance d, for both $H = H_0$ and $H = H_1$. This is shown in Figure 4.3, where the true state of the channel is $H = H_1$, and the *a posteriori* distribution of the distance, conditioned on the values of the available estimates for both LOS and NLOS scenarios is shown. In this example, for the known value of d, the likelihood that the available estimates occurred due to NLOS propagation rather than LOS propagation is higher, and the maximum-likelihood (ML) channel state [34] would be $\hat{H} = H_1$. Evidently, if d is known, based on the likelihood values of *any* of the given estimates, the ML channel state can be directly obtained from the known conditional distributions of the estimates.

In reality however, we do not know the physical distance d between the nodes. Indeed, for localization applications, ascertaining the distance d accurately is a primary objective, and



(a) (b)

Figure 4.3: *a posteriori* density function of the distance conditioned on (a) the TOA estimate x and the channel state H, and (b) the RSS estimate y and the channel state H. For both plots, the true channel state is $H = H_1$.

is typically done using TOA or RSS estimates. For simplicity, we assume that the distance is estimated using the TOA:

$$\hat{d} = cx_1.$$

While the above distance estimate is unbiased when $H = H_0$, it is positively biased with high probability when $H = H_1$, and therefore, the likelihood of the TOA estimate cannot be used to identify the channel state as seen in Figure 4.3. However, for the given RSS estimate x_2 when $H = H_1$, the probability that the received RSS estimate was obtained under NLOS propagation conditions, even with the biased distance estimate $\hat{d}$, is higher than the corresponding LOS probability. This suggests that joint NLOS identification based on all three estimates can improve the probability of successful identification, even with noisy and/or biased distance estimates.

The formulation of the joint channel state identification is as follows: our initial hypothesis is that the state of the channel is LOS, i.e, $\hat{H} = H_0$, and that $\hat{d} = cx_1$. For each of the given estimates, we determine the conditional probabilities:

$$\begin{aligned} p_i &= \Pr\left\{X_i = x_i \middle| \hat{d},\ H_0\right\}, \\ q_i &= \Pr\left\{X_i = x_i \middle| \hat{d},\ H_1\right\}. \end{aligned}$$

We then compute $D_L = \prod_i p_i$ and $D_N = \prod_i q_i$. It must be pointed out that if the estimates are assumed to independent for a fixed distance and channel state, D_L and D_N are the joint conditional probabilities of the estimates. The values of D_L and D_N can then serve as the

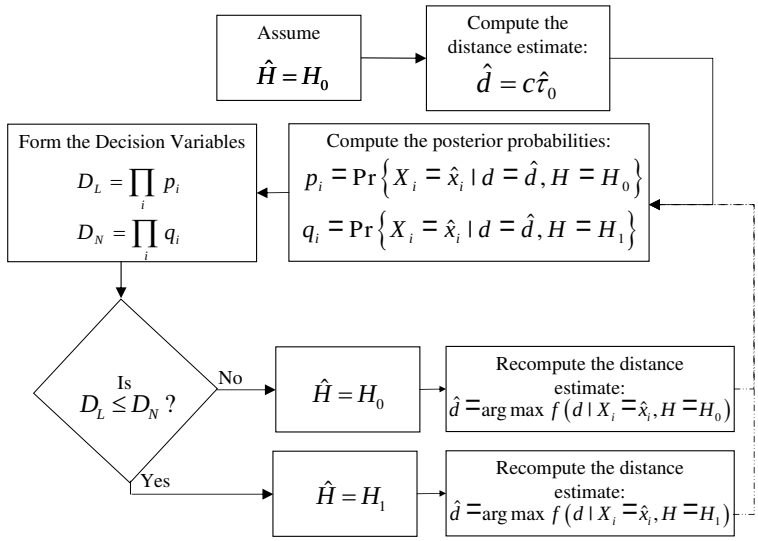


Figure 4.4: Block Diagram of statistical decision-theoretic framework: the joint probabilities D_L and D_N are compared to determine the channel state estimate.

decision metrics for estimating the channel state. In particular, if $D_N < D_L$, we change our hypothesis of the channel state from LOS to NLOS:

$$\hat{H} = \begin{cases} H_0, & D_L > D_N, \\ H_1, & D_L < D_N. \end{cases} \tag{4.6}$$

After the state of the channel has been determined, the range estimate can then be updated:

$$\hat{d} = \arg\max_d \ \Pr\left\{ X_i = x_i \,\middle|\, d,\ \hat{H} \right\}, \tag{4.7}$$

which is the ML estimate of d for the given estimate x_i and the estimated channel state $\hat{H}$. Once the channel state and the new distance estimate have been computed, these can be fed back to verify if the conditional probabilities change, as shown in Figure 4.4. The above method can be viewed as joint iterative ML estimation of the channel state, with the distance d treated as a nuisance parameter. As seen in Figure 4.4, the distance d is iteratively estimated along with the channel state. The iteration is terminated when the estimated channel state does not change with additional iterations. It has been observed that the above iterative scheme typically converges within two iterations.

Clearly, there are variations to this scheme: we may choose to use only a subset of the estimates, depending upon their availability and/or knowledge of their statistical distributions. Specifically, in Figure 4.4, to determine the decision metrics D_L and D_N, we may choose to use only TOA, or RSS, or RMS conditional probabilities, or any combination of these estimates.

In the following section, we provide simulation results that compare the efficacy of combinations of different estimates in distinguishing between LOS and NLOS propagation. We show that, based on the indoor UWB measurements in [1], the RDS estimates serve as efficient "differentiators" between LOS and NLOS propagation scenarios. We further show that even in the absence of information on their statistical dependence on distance, the RDS estimates can be used through conventional hypothesis-testing to accurately identify the channel state.

4.4 Simulation Results

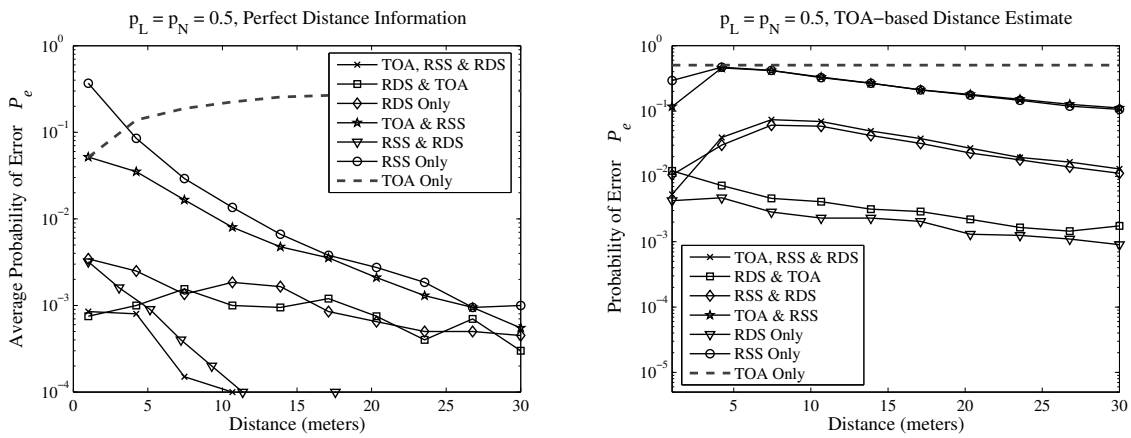
The performance of decision-theoretic NLOS identification, based on the joint statistics of combinations of TOA, RSS and RDS estimates, is shown in Figures 4.5(a) and 4.5(b). There exist 7 combinations of these estimates: "Only-TOA", "Only-RSS", "Only-RDS", "TOA & RSS", "RSS & RDS", "RDS & TOA" and "TOA, RSS & RDS". A large number of realizations of the TOA, RSS and RDS estimates are generated for each value of d between 1 and 30 meters using (4.1), (4.2) and (4.5). A given realization of an estimate is generated based on its corresponding conditional distribution for the LOS scenario ($H = H_0$) with probability p_L, and using the conditional distribution for the NLOS scenario ($H = H_1$) with probability $p_N = 1 - p_L$. The described hypothesis-testing approach is then applied to the each realization of the given parameters, and the channel state $\hat{H}$ is estimated. For each value of the distance, the average probability of channel state estimation error P_e is evaluated based on the relation:

$$P_e = p_L \cdot \Pr\left\{\hat{H} = H_1 \middle| H = H_0\right\} + p_N \cdot \Pr\left\{\hat{H} = H_0 \middle| H = H_1\right\}. \tag{4.8}$$

In the absence of the knowledge of *a priori* probabilities, the channel states are assumed to be equi-probable: $p_L = \Pr\{H = H_0\} = \Pr\{H = H_1\} = p_N = 0.5$, and 10,000 realizations of the TOA, RSS, and the RDS were generated for each value of d.

Figure 4.5(a) compares the average probability of channel state estimation error P_e, when the distance between the transmit and receive nodes is known exactly, and the TOA, RSS, and RDS estimates obey the statistical distributions with the parameters given in Table 4.1, defined in (4.1), (4.2) and (4.5) respectively. Given different combinations of the estimated parameters, the ML estimate of the channel state can be obtained as discussed in Section 4.3. Although exact distance information is typically not available, the performance of the channel state identification with perfect distance information (i) serves as an upper bound for the probability of successful identification in the case where the distance d is unknown, and (ii) helps quantify the sensitivity of the discussed channel identification techniques to the accuracy of distance estimates.

Figure 4.5(b) compares the values of P_e obtained using different combinations of the estimated parameters when d is unknown and is estimated using the iterative procedure discussed in section 4.3. It must be pointed out that in all these cases, the initial distance



(a) (b)

Figure 4.5: Probability of Error using different methods when (a) d is known exactly, and (b) d is estimated using the TOA estimate. In the above simulations, it is assumed that the probability of LOS propagation is equal to the probability of NLOS propagation.

estimate $\hat{d}$ is obtained using the TOA estimate $\hat{\tau}_0$. It is intuitive that the "only-TOA" method produces approximately 50% successful identification in Figure 4.5(b), as the initial hypothesis is always $\hat{H} = H_0$ and this, with high probability, results in $p_1 > q_1$. From Figures 4.5(a) and 4.5(b), we observe that while the method which uses all three estimates ("TOA, RSS & RDS") results in the lowest value of P_e of all the methods when perfect distance information is available (Figure 4.5(a)), this is not the case when perfect distance information is not available (Figure 4.5(b)). We further note that as the distance d increases, the average probability of channel state estimation error decreases, as the regions of overlap for the density functions in LOS and NLOS scenarios diminish, leading to more accurate channel state identification. This results from the fact that the mean of the estimates grows at different rates in LOS and NLOS scenarios, and the difference of the mean values in the two scenarios increases as d increases.

A key observation in both Figures 4.5(a) and 4.5(b) is that methods incorporating the RDS estimate are most successful in differentiating between the LOS and NLOS channel states. Even in isolation, the RDS estimate can be used to distinguish between LOS and NLOS propagation conditions with a high probability of success, and adding TOA and RSS statistical information does not result in considerable gains when accurate distance information is not available. Indeed, for the statistical parameters observed in [1] and listed in Table 4.1, in the absence of perfect distance information, using TOA and RSS estimates in addition to the RDS estimates actually degrades the probability of successful channel state estimation, as observed in Figure 4.5(b).

4.4.1 Absence of Complete Statistical Information

A major issue concerning the practical application of the statistical decision-theoretic method described in section 4.3 is the availability of complete statistical information corresponding to received signal parameters, and particularly their variation with distance. When the conditional distributions of the received signal parameters are unknown, we may resort to "conventional" hypothesis-testing, where the received signal parameter is compared with a threshold, followed by a decision on the channel state depending on whether the estimate is smaller or larger than the threshold. Due to the strong dependence of the statistics of the TOA and RSS estimates on the true distance between transmit and receive antennas, the threshold used for hypothesis-testing these estimates must be dependent on the true distance. However, as seen in Figure 4.5(b), the lack of perfect distance information results in considerable performance degradation when using TOA and RSS estimates with noisy and/or biased distance estimates.

On the other hand, we observed in Figure 4.5(b) that significantly lower channel state estimation error probabilities can be obtained with RDS estimates, even with noisy and biased distance estimates. Figure 4.6 shows the histogram of the RDS estimates in LOS and NLOS scenarios, *for all measurement distances* (1-30 meters) [1]. We see that regardless of d, the regions of support for the density functions of the RDS estimates remain sufficiently separated to use conventional hypothesis testing, without requiring a distance-dependent threshold. The statistics of the RDS estimates observed in several measurement campaigns [59], [2], summarized in Table 4.2, indicate that the distinction between the RDS estimates in LOS and NLOS propagation scenarios is general, and not restricted to a given set of measurements. The statistics of the RDS estimates corresponding to the channel model adopted by the IEEE 802.15.3a subcommittee [63] are also shown. It is important to emphasize that for each of the measurement scenarios listed in Table 4.2, differences in the propagation environments, deconvolution algorithms and their associated parameters, types and directionality of antennas, etc., result in the variation of the observed values of the RDS across the measurement campaigns. We would like to (a) investigate the performance of conventional hypothesis-testing of the RDS estimates for the sets of measurements listed in Table 4.2, and (b) quantify the extent of performance degradation due to the lack of the knowledge on the statistical dependence on distance.

4.4.2 Definition of a distance metric

Assuming that the RDS estimate is positive and Gaussian [59], as given in (4.5), the minimum probability of error $P_e^\star$ and the optimal threshold $T^\star$ for hypothesis-testing can be computed in a straightforward manner, as detailed in Appendix 4B. In general, the values of $P_e^\star$ and $T^\star$, respectively given by (4.39) and (4.40), depend on the means $\{\mu_L, \mu_N\}$ and the standard deviations $\{\sigma_L, \sigma_N\}$ of the RDS estimates. In order to allow for a unified comparison of

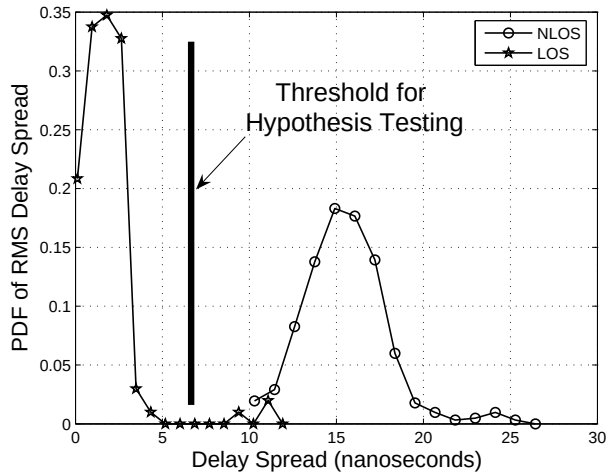


Figure 4.6: Hypothesis testing of the RMS delay spread for the nature of the UWB channel: the histogram of the RMS delay spread in LOS and NLOS scenarios for distances 1-30 meters. We see that since the PDF of the RMS delay spread for the LOS and NLOS cases are well separated, the performance of NLOS identification using this approach is not sensitive to the value of the threshold selected.

performance using different sets of measurements, we define a metric:

$$\mathcal{M} = \frac{\sqrt{\sigma_L \sigma_N}}{|\mu_N - \mu_L|}. \tag{4.9}$$

Intuitively, the metric $\mathcal{M}$ is a measure of extent of the overlap of the probability density functions of the RDS estimates in LOS and NLOS scenarios: as $\mathcal{M}$ decreases, $P_e^\star$ decreases. As the parameters $\{\mu_L, \mu_N, \sigma_L, \sigma_N\}$ vary, the values of the metric vary as well. However, as seen in Figure 4.7, there is a strong correlation between the value of $\mathcal{M}$ defined above and the value of $P_e^\star$ computed using (4.39). Further, we see that the value of $P_e^\star$ achieved largely depends on the value of $\mathcal{M}$, rather than the specific values of the parameters $\{\mu_L, \mu_N, \sigma_L, \sigma_N\}$.

Table 4.3 lists the values of $\mathcal{M}$, $P_e^\star$ and $T^\star$ respectively computed using (4.9), (4.39) and (4.40) from the RDS statistics in Table 4.2. For each of these measurement statistics (S1-S9 in Table 4.2), the values of $P_e^\star$ are plotted versus the values of $\mathcal{M}$ in Figure 4.7. We see that for the sets of measurements S1-S9, the probability of channel state estimation error can vary from 10^{-5} to 0.2. It must be emphasized that these values of P_e are obtained via simulations, under the assumption that the distribution of the RDS estimate under LOS and NLOS conditions is positive and Gaussian distributed. The direct application of the hypothesis-testing to received signal measurements over all distances [1] results in a probability of error P_e of approximately 0.02 (see (M1) in Figure 4.7). With (M1), no knowledge of the distance or statistical distributions was assumed and the RDS estimates from received signal measurements were compared to a threshold $T = 9.5$ nanoseconds. This value of T was

Table 4.2:]
RMS Delay Spread Statistics from Measurement Campaigns [2]

Indoor LOS		
Measurement Campaign	Mean μ_L (nanoseconds)	Standard Deviation σ_L (nanoseconds)
AT&T [62], [64]	4.71 3.55	2.31 1.65
Samsung/SAIT [65]	14.00 12.87	1.53 1.87
Time Domain [61]	5.27	3.37
MPRG, VT [1]	3.34	2.17
IEEE CM1 [66]	6	1.1
Indoor NLOS		
Measurement Campaign	Mean μ_N (nanoseconds)	Standard Deviation σ_N (nanoseconds)
AT&T [62], [64]	8.20 7.35	3.30 3.45
Samsung/SAIT [65]	38.61 26.51	8.03 5.22
Time Domain [61]	14.59	3.41
MPRG, VT [1]	16.08	2.41
IEEE CM2 [66]	8	0.75
IEEE CM3 [66]	14.5	2.26
IEEE CM4 [66]	25	3.7

Table 4.3: Computation of $\mathcal{M}$, $P_e^\star$ and $T^\star$ using the statistics in Table 4.2

Metric Calculations			
Measurement Campaign	Metric $\mathcal{M}$	Minimum Probability of Error $P_e^\star$	Optimal Threshold $T^\star$ (nanoseconds)
AT&T (S1, S2)	0.50	0.16	6.69
	0.63	0.20	5.69
Samsung/SAIT (S3, S4)	0.14	3.5×10^{-3}	18.35
	0.23	0.024	17.1
Time Domain (S5)	0.36	0.087	10.1
MPRG, VT [1] (S6)	0.18	3.3×10^{-3}	9.49
IEEE CM1-CM2 (S7)	0.43	0.13	6.97
IEEE CM1-CM3 (S8)	0.19	5.5×10^{-3}	8.86
IEEE CM1-CM4 (S9)	0.10	3.35×10^{-5}	10.76

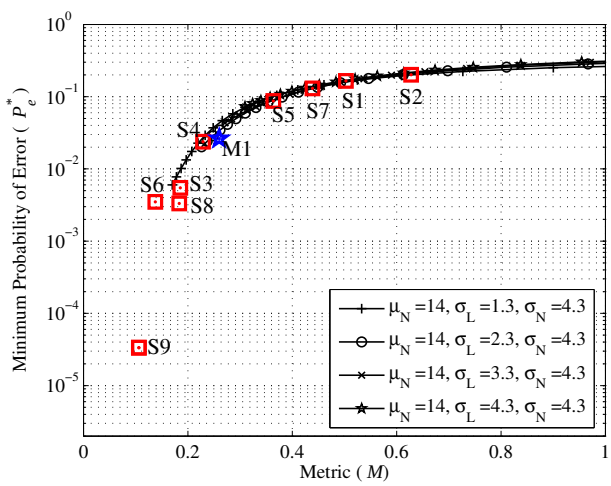


Figure 4.7: The minimum probability of error $P_e^\star$ versus the metric $\mathcal{M}$ for different values of $\{\mu_L, \sigma_L, \mu_N, \sigma_N\}$. The value of μ_L is varied to obtain different values of the metric. Simulated values of $P_e^\star$ from Table 4.3 (S1-S9) are also shown.

computed using the analysis in Appendix 4B, and was found to provide robust performance over a large number of measurement locations. The average probability of error in scenario (M1) is about an order of magnitude larger than the "Only-RDS" scenario discussed in section 4.3 equipped with distance-dependent statistical knowledge, where P_e averaged over d was approximately 2×10^{-3}. While this suggests that the lack of knowledge on the statistical dependence on distance does result in performance degradation, the performance may be acceptable for localization applications. In Chapter 6, we present measurement results from a location-tracking experiment that demonstrate that hypothesis-testing of the RDS estimate provides accurate channel identification from the perspective of localization accuracy.

4.5 Conclusions from Part I

In part I, we presented a statistical decision-theoretic approach to NLOS identification based on TOA, RSS and RDS estimates. Based on the parameters from extensive measurement results, we characterized the ability of TOA, RSS and RMS delay spread estimates to distinguish between LOS and NLOS propagation. We showed that while utilizing all three estimates jointly in detecting the nature of propagation yields the best results in the presence of accurate distance information, the RMS delay spread estimate is the best indicator of the nature of propagation when perfect distance information is not available. Our results suggest that hypothesis-testing of the RMS delay spread estimates of UWB signals can serve as a robust and computationally-efficient means of identifying NLOS propagation, and does not require distance information or considerable statistical information. The efficacy of hypothesis testing of the RMS delay spread will be demonstrated in a location-tracking experiment with indoor UWB measurements in Chapter 6.

Part II: NLOS mitigation using Linear Programming

4.6 Impact of NLOS bias errors

4.6.1 Notation, Models and Assumptions

Suppose that an "unlocalized" node's (unknown) location is $\boldsymbol{x} = [x\ y]^T$. Let $\mathcal{A}$ denote the set of anchor locations, and $\mathcal{L} \in \mathcal{A}$ denote the set of locations of anchors which provide LOS range estimates, with cardinality $m_L = |\mathcal{L}|$. The known locations of the LOS anchors are denoted by $\{\boldsymbol{x}_{Lj}\}$, $j = 1, 2, \cdots, m_L$. Similarly, $\mathcal{N} \in \mathcal{A}$ represents the set of locations of anchors that provide NLOS range estimates, with $m_N = |\mathcal{N}|$, and the known set of locations

of the NLOS anchors is represented by $\{\boldsymbol{x}_{Nj}\}$, $j = 1, 2, \cdots, m_N$.

The LOS range estimates $\{r_{Lj}\}$, $j = 1, 2, \cdots, m_L$, are modeled as unbiased Gaussian [9] estimates of the actual inter-node distances $R_{Lj} = \|\boldsymbol{x} - \boldsymbol{x}_{Lj}\|$:

$$r_{Lj} = R_{Lj} + n_{Lj}, \quad j = 1, 2, \cdots, m_L, \tag{4.10}$$

where n_{Lj} represents zero-mean Gaussian range measurement noise in the jth LOS range estimate: $n_{Lj} \sim \mathcal{N}(0, \sigma_{Lj}^2)$. The range measurement noise variance σ_{Lj}^2 can be modeled as [67]:

$$\sigma_{Lj}^2 = K_E R_{Lj}^{\beta_L}, \tag{4.11}$$

where β_L is the LOS path loss exponent and K_E is a proportionality constant (governed by the transmit power and the receiver noise floor) that determines the accuracy of range estimation. This model arises due to the fact that the accuracy of TOA-based range estimates can be shown [17], [36] to be inversely proportional to the received signal-to-noise ratio (or, more generally, the signal-to-interference-and-noise ratio) assuming matched-filter detection. The above model for the accuracy of range estimates [67] applies to both TOA and RSS-based range estimates when $\beta = 2$. The vector of LOS range estimates is denoted by $\boldsymbol{r_L} = [r_{L1}\ r_{L2}\ \cdots\ r_{Lm_L}]^T_{(1\times m_L)}$, where $[\cdot]_{(1\times n)}$ denotes a row vector of length n.

The NLOS range estimates are assumed to be positively biased Gaussian estimates [49] of the true distances:

$$r_{Nj} = R_{Nj} + n_{Nj} + b_{Nj}, \quad j = 1, 2, \cdots, m_N, \tag{4.12}$$

where $R_{Nj} = \|\boldsymbol{x} - \boldsymbol{x}_{Nj}\|$, $n_{Nj} \sim \mathcal{N}(0, \sigma_{Nj}^2)$, $\sigma_{Nj}^2 = K_E R_{Nj}^{\beta_N}$, b_{Nj} are the NLOS bias errors, and β_N is the effective path loss exponent under NLOS conditions. We assume that the bias errors are always positive: $b_{Nj} > 0, \forall j$. The bias errors b_{Nj} and the range measurement noise n_{Nj} are assumed to be independent random variables. Although we make no assumptions about the statistical distribution [53] of the NLOS bias errors in the following development, for the purpose of simulation, we assume that bias errors are uniformly distributed: $b_{Nj} \sim \mathcal{U}(0, B_{\max})$, where $B_{\max}$ represents the maximum possible bias[1]. Additionally, we assume that the bias errors are, with high probability, much larger than the range measurement noise $B_{\max} \gg \sigma_{Nj}, \quad j = 1, 2, \cdots, m_N$. Finally, without loss of generality, we assume that the coordinate axes are selected such that $\boldsymbol{x} \geq \boldsymbol{0}$.

4.6.2 Discarding NLOS range estimates

As we shall show in a later section, when we have at least three range estimates, the LS estimator [20] can be used to compute an estimate $\hat{\boldsymbol{x}}$ of the unlocalized node's location $\boldsymbol{x}$. We define the *localization error*, a measure of the accuracy of the location-estimate $\hat{\boldsymbol{x}}$, as:

$$\Omega = \|\boldsymbol{x} - \hat{\boldsymbol{x}}\|^2 \quad (\text{meter}^2). \tag{4.13}$$

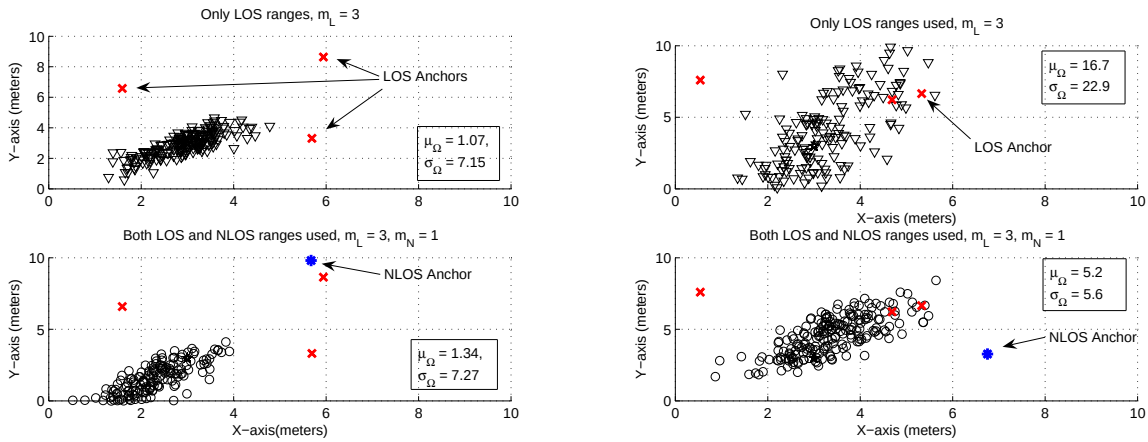
[1]Practically speaking, $B_{\max}$ would depend on the propagation environment and physical layer parameters such as the transmit power.

It must be noted that Ω is a random variable, with different instances corresponding to different realizations of the range measurement noise, bias errors and anchor locations. Therefore, we characterize the accuracy of location-estimates through the mean μ_Ω and standard deviation σ_Ω of the localization error defined in (4.13); smaller values of both μ_Ω and σ_Ω indicate more accurate node location-estimates on the average.

When $m_N = 0$ and $m_L \geq 3$, the LS estimator provides accurate estimates of a node's location [20]. However, when $m_N > 0$, we need effective ways of incorporating NLOS information into the estimation procedure. The Cramer-Rao Lower Bound (CRLB) analysis presented in [52], [53] showed that, in the absence of prior statistical information on the NLOS range estimates, the MVUE discards the biased NLOS range estimates and utilizes only LOS range information while computing location estimates. However, this approach may not be optimal when using practical estimators, such as the LS estimator, that do not achieve the CRLB. Figures 4.8(a) and 4.8(b) show the impact of directly (without mitigation of the bias errors) incorporating NLOS range estimates into the LS solution for two specific scenarios. For the specific distribution of anchors shown in Figure 4.8(a), directly incorporating the NLOS ranges into the LS solution without any mitigation of the bias in the range estimate can degrade average localization accuracy defined in terms of μ_Ω and σ_Ω. However, in some cases, and in particular for the example shown in Figure 4.8(b), introducing the NLOS range estimate directly into LS location estimation can *improve* performance in terms of μ_Ω and σ_Ω.

Generally speaking, it is observed that discarding the NLOS range estimates does not result in poor performance when the geometry of LOS anchor nodes has certain properties, best described by the *geometric dilution of precision* (GDOP) [9], where a large GDOP (as defined in [9]) implies poor localization accuracy. It has been observed that when the GDOP of LOS anchors is large, the presence of an additional NLOS range estimate results in an improvement in performance: the addition of a NLOS node typically reduces the effective GDOP and this compensates for the inaccuracy of the NLOS range estimate.

These two examples show that (i) directly incorporating NLOS range estimates into existing practical estimators without reducing the impact of bias errors can adversely affect localization accuracy; however, (ii) we do not wish to discard the NLOS range estimates, since their use could improve the performance of practical estimators under certain conditions. Indeed, in indoor networks, we may have more NLOS range estimates than LOS range estimates. Therefore, what is desired is a method that allows the "soft-activation" of NLOS range information: the NLOS range estimates are not incorporated directly, but are used in conjunction with LOS range estimates when LOS range estimates alone do not guarantee accurate node location estimates. In the following sections, a LP approach that achieves this goal is described.



(a) (b)

Figure 4.8: **(a)** This example shows several instances of the LS location estimate $\hat{\boldsymbol{x}}$, one for each realization of the range estimates, for (i) (top) Only $m_L = 3$ LOS estimates (ii) (bottom) including $m_L = 3$ LOS and $m_N = 1$ NLOS range estimates. The NLOS range estimate is treated exactly like an LOS range estimate and directly incorporated into the LS solution in the bottom figure. In this case, the addition of the biased NLOS range estimate *degrades* localization accuracy with respect to μ_Ω and σ_Ω. **(b)** In this case, the addition of the biased NLOS range estimate *improves* localization accuracy with respect to μ_Ω and σ_Ω. In both cases, $\boldsymbol{x} = [3\ 3]^T$, $\beta_L = 2$, $\beta_N = 2.5$, $K_E = 0.1$, $B_{\max} = 4$ meters.

4.7 A Linear Programming Approach

In this section, we show that the problem of node localization given LOS range information can be cast into the form of a linear program. We then modify the linear program to utilize additional NLOS range information, resulting in a method that utilizes a mixture of LOS and NLOS range estimates to estimate a node's location.

4.7.1 LOS range estimates

The LOS range estimates, which are modeled as unbiased estimates of the true ranges, can be used to define conditions satisfied by the unknown node location $\boldsymbol{x}$. We can write:

$$\|\boldsymbol{x} - \boldsymbol{x}_{Li}\| = r_{Li} \Rightarrow (x - x_{Li})^2 + (y - y_{Li})^2 = r_{Li}^2, \quad i = 1, 2, \cdots, m_L. \tag{4.14}$$

These relations are non-linear equations in x and y and represent the fact that $\boldsymbol{x}$ lies on a circle of radius r_{Li} whose center is $\boldsymbol{x}_{Li}$. This system of equations can be linearized by extracting the difference of each of these equations from the others, forming a linear system of $M = \binom{m_L}{2}$ *distinct* equations:

$$\begin{aligned} a_{ij}x + b_{ij}y = c_{ij}, \quad i,j &= 1, 2, 3, \cdots, m_L, \ i < j \\ \text{where } a_{ij} = x_{Li} - x_{Lj}, \ b_{ij} &= y_{Li} - y_{Lj}, \ c_{ij} = \frac{\left(x_{Li}^2 - x_{Lj}^2\right) + \left(y_{Li}^2 - y_{Lj}^2\right) - \left(r_{Li}^2 - r_{Lj}^2\right)}{2}. \end{aligned} \tag{4.15}$$

Each of these M equations can be viewed as representing the lines connecting the intersection points (if any) of pairs of circular constraints defined in (4.14). As the range estimates defined in (4.10) are noisy, in general, $r_{Li} \neq R_{Li}$, and solving these equations simultaneously may not yield a unique solution. Resorting to an error minimization approach, for every potential solution $\boldsymbol{x}$, and for every equation, we can define the residual error as:

$$e_{ij} = a_{ij}x + b_{ij}y - c_{ij}, \quad i, j = 1, 2, 3, \cdots, m_L, \ i < j. \tag{4.16}$$

The final estimate $\hat{\boldsymbol{x}}$ can be selected such that an objective function Z, such as the sum of the residual error squares, is minimized:

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} Z = \arg\min_{\boldsymbol{x}} \sum_i \sum_{j,\ j>i} e_{ij}^2.$$

This is the equivalent to the LS approach defined in [20]. It is important to note that (i) although the system of equations in (4.15) is linear, the objective function Z is non-linear in x and y, and (ii) we require $m_L \geq 3$ to form an unambiguous solution[2]. The LS solution is given by

$$\hat{\boldsymbol{x}} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\boldsymbol{c}, \tag{4.17}$$

[2]In the three-dimensional case, we require $m_L \geq 4$.

where

$$\boldsymbol{A} = \begin{bmatrix} a_{12} & a_{13} & \cdots & a_{1m_L} & a_{23} & \cdots & a_{(m_L-1)m_L} \\ b_{12} & b_{13} & \cdots & b_{1m_L} & b_{23} & \cdots & b_{(m_L-1)m_L} \end{bmatrix}_{2\times M}^T, \tag{4.18}$$

and

$$\boldsymbol{c} = \begin{bmatrix} c_{12} & c_{13} & \cdots & c_{1m_L} & c_{23} & \cdots & c_{(m_L-1)m_L} \end{bmatrix}_{1\times M}^T. \tag{4.19}$$

Looking at (4.16), we see that the set of variables $\{e_{ij}\}$ plays the role of unconstrained "slack variables" [68] in the system of M equations. Therefore, this linear system of equations can be converted to a linear program [68] if the objective function Z is a linear function of the unknowns. Specifically, if we define

$$Z \triangleq \sum_i \sum_{j,\ j>i} |e_{ij}|,$$

and then replace the unconstrained variable e_{ij} by $e_{ij}^+ - e_{ij}^-$, $e_{ij}^+ \geq 0,\ e_{ij}^- \geq 0$, we can write an alternative *linearized* objective function that is to be minimized, as

$$Z \triangleq \sum_i \sum_{j,\ j>i} \left(e_{ij}^+ + e_{ij}^-\right). \tag{4.20}$$

It must be noted that in the optimal solution that minimizes Z, only one term among $\{e_{ij}^+, e_{ij}^-\}$ will be equal to $|e_{ij}|$, with the other being zero [68]. The constraints are then given by

$$a_{ij}x + b_{ij}y - e_{ij}^+ + e_{ij}^- = c_{ij}, \qquad i,j = 1,2,3,\cdots,m_L,\ i<j. \tag{4.21}$$

Since there are now $2M$ non-negative slack variables, the vector $\boldsymbol{z}$ of $(2M+2)$ variables can be written as $\boldsymbol{z} = \left[x\ y\ \boldsymbol{\epsilon}^T\right]^T$, where

$$\boldsymbol{\epsilon} = \left[e_{12}^+\ e_{12}^-\ e_{13}^+\ e_{13}^-\ \cdots\ e_{(m_L-1)m_L}^+\ e_{(m_L-1)m_L}^-\right]_{1\times 2M}^T. \tag{4.22}$$

Thus, the linear program can be formulated in *standard form* [68] as

$$\begin{aligned} \min\ Z &= \boldsymbol{f}_L^T \boldsymbol{z}, \quad \text{such that} \\ [\boldsymbol{A} \mid \boldsymbol{J}]\,\boldsymbol{z} &= \boldsymbol{c},\ \boldsymbol{z} \geq \boldsymbol{0}, \end{aligned} \tag{4.23}$$

where $\boldsymbol{A}$ and $\boldsymbol{c}$ were respectively defined in (4.18) and (4.19),

$$\boldsymbol{J} = \begin{bmatrix} -1 & 1 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & \cdots & \cdots & 0 & 0 \\ \vdots & \vdots & & & & & & \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix}_{M\times 2M}, \tag{4.24}$$

and $\boldsymbol{f}_L = [\boldsymbol{0}_{2\times 1}^T\ \boldsymbol{1}_{2M\times 1}^T]^T$. Here, $\boldsymbol{0}_{k\times l}$ represents a $k\times l$ matrix of zeros and $\boldsymbol{1}_{k\times l}$ represents a $k\times l$ matrix of ones.

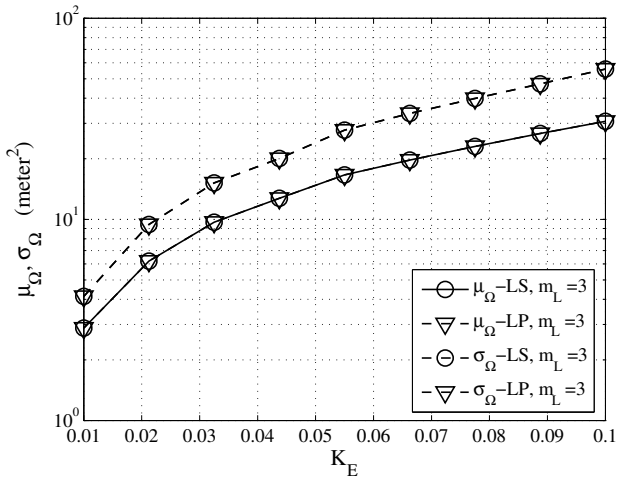


Figure 4.9: Standard Deviation of the Localization Error Ω, $m_L = 3$, $m_N = 0$ for a specific distribution of anchors.

It can be seen from Figure 4.9 that the linearization of the objective function does not result in significant degradation of the localization accuracy. Therefore, we now have a linear program that can be used to solve for a node's location given LOS range estimates. In this linear program, the objective function Z defined in (4.20) is a function of the distances of a point $\boldsymbol{x}$ to the straight lines given in (4.15). If we use NLOS range estimates in a similar manner, by incorporating them into the objective function we could potentially degrade the accuracy of the location estimate. Instead, as described in the following section, we can use the NLOS range estimates to constrain the feasible region for $\boldsymbol{x}$ without affecting the objective function defined using LOS range estimates, thereby limiting the possibility of large errors, particularly when the number of LOS range estimates is small.

4.7.2 NLOS range estimates

As the bias errors in the NLOS range estimates are always positive, and are assumed to be much larger than the range measurement noise, we know each NLOS range estimate r_{Ni} defined in (4.12) is, with high probability, larger than the true range R_{Ni}, $i = 1, 2, \cdots, m_N$. Based on this observation, we can convert the NLOS range estimates into inequalities for $i = 1, 2, \cdots, m_N$:

$$\|\boldsymbol{x} - \boldsymbol{x}_{Ni}\| \leq r_{Ni} \Rightarrow (x - x_{Ni})^2 + (y - y_{Ni})^2 \leq r_{Ni}^2. \tag{4.25}$$

These inequalities imply that the feasible region for $\boldsymbol{x}$ lies *in the interior* of each of the circular constraints defined by (4.25). Note that this assumption cannot be made if the standard deviation of the zero-mean measurement noise and the positive bias in (4.12) are comparable.

Once again, these are non-linear constraints on x and y. However, these constraints can be relaxed to the following linear constraints, as suggested in [56]:

$$x - x_{Ni} \leq r_{Ni},\ -x + x_{Ni} \leq r_{Ni},\ y - y_{Ni} \leq r_{Ni},\ -y + y_{Ni} \leq r_{Ni}, \quad (i = 1, 2, \cdots, m_N). \tag{4.26}$$

This essentially relaxes the circular constraints to rectangular constraints as shown in Figure 4.10(a). It is readily seen that the new rectangular feasible region contains the original (convex) feasible region formed by the intersection of the original circular regions. We can now write the above four constraints for the ith NLOS range estimate in standard form [68]:

$$\begin{aligned} &x - x_{Ni} + u_{1i} = r_{Ni},\ -x + x_{Ni} + u_{2i} = r_{Ni},\ y - y_{Ni} + v_{1i} = r_{Ni},\ -y + y_{Ni} + v_{2i} = r_{Ni} \\ &u_{1i},\ u_{2i},\ v_{1i},\ v_{2i} \geq 0, \qquad i = 1, 2, \cdots, m_N. \end{aligned} \tag{4.27}$$

where $u_{1i},\ u_{2i},\ v_{1i},\ v_{2i}$ are the slack variables corresponding to the ith NLOS range estimate. Defining $\boldsymbol{w}_i = [u_{i1}\ u_{2i}\ v_{i1}\ v_{2i}]^T_{1\times 4}$ and $\boldsymbol{z}_i = [x\ y\ \boldsymbol{w}_i^T]^T$ as the vectors of variables corresponding to the ith NLOS range estimate, we can express the above equations in matrix form as

$$[\boldsymbol{B}_1 |\ \boldsymbol{I}_{4\times 4}]\, \boldsymbol{z}_i = \boldsymbol{r}_i,\ \ \boldsymbol{z}_i \geq \boldsymbol{0},$$

where

$$\boldsymbol{B}_1 = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{bmatrix},\ \boldsymbol{r}_i = \begin{bmatrix} r_{Ni} + x_{Ni} \\ r_{Ni} - x_{Ni} \\ r_{Ni} + y_{Ni} \\ r_{Ni} - y_{Ni} \end{bmatrix},$$

and $\boldsymbol{I}_{n\times n}$ denotes an $n \times n$ identity matrix. We can now stack the constraints corresponding to each of the m_N NLOS range estimates to form a system of $N = 4m_N$ equations as follows:

$$[\boldsymbol{B}\ \boldsymbol{I}_{N\times N}]\, \boldsymbol{z} = \boldsymbol{r},\ \ \boldsymbol{z} \geq \boldsymbol{0},$$

where

$$\boldsymbol{B} = \begin{bmatrix} \boldsymbol{B}_1^T & \boldsymbol{B}_1^T & \cdots & \boldsymbol{B}_1^T \end{bmatrix}^T_{(2\times N)},\ \boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_1^T & \boldsymbol{r}_2^T & \cdots & \boldsymbol{r}_{m_N}^T \end{bmatrix}^T_{(1\times N)}, \tag{4.28}$$

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1^T & \boldsymbol{w}_2^T & \cdots & \boldsymbol{w}_{m_N}^T \end{bmatrix}^T_{(1\times N)}, \tag{4.29}$$

with the vector of variables being defined as

$$\boldsymbol{z} = \begin{bmatrix} x & y & \boldsymbol{w}^T \end{bmatrix}^T_{(1\times(N+2))}.$$

It is important to note that in the above analysis, no objective function was defined based on the NLOS range estimates, and only a feasible region for $\boldsymbol{x}$ was derived. The feasible region can further be constrained by including the tangents at the intersection points of the circular constraints defined in (4.25) to reduce the size of the feasible region as shown in Figure 4.10(b). In the following subsection, we integrate the constraints and objective function obtained using LOS range estimates with the NLOS constraints defined above, for the problem of node location-estimation given any mixture of LOS and NLOS range estimates such that $m_L \geq 3$, $m_N \geq 0$.

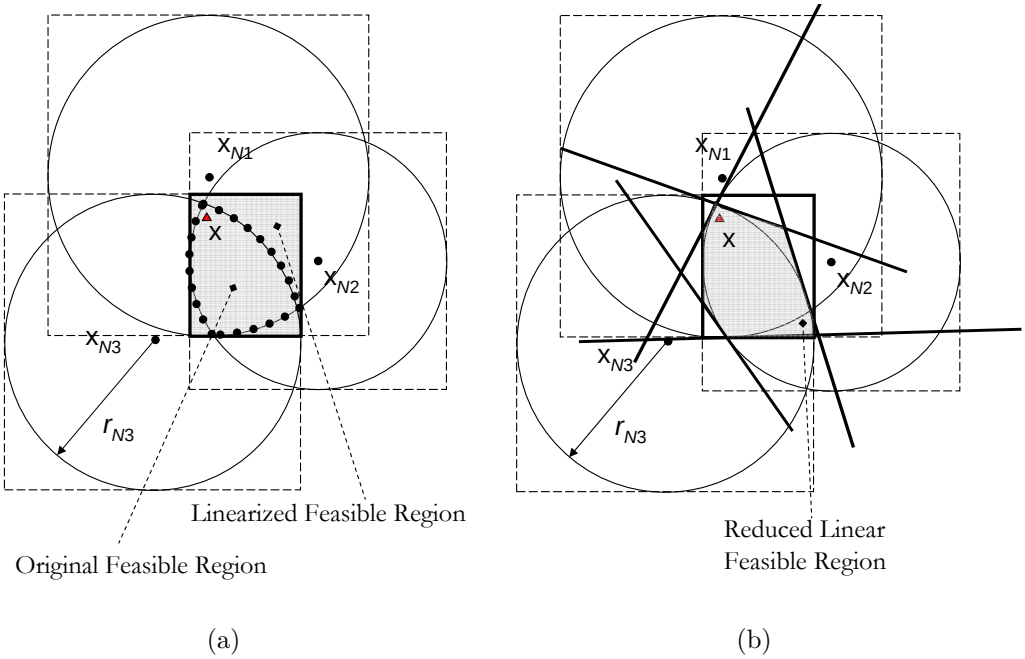


Figure 4.10: (a) Linearization of $m_N = 3$ NLOS constraints: the NLOS circular constraints are converted to rectangular constraints. (b) The tangents at the intersections can be used to further reduce the size of the feasible region.

4.7.3 Combining the LOS and NLOS Range Information

Based on the above subsections, given $m_L \geq 3$ LOS range estimates and $m_N \geq 0$ NLOS range estimates, we can combine them into a single linear program. We define the vector of variables as

$$\boldsymbol{z} = [x \;\; y \;\; \boldsymbol{\epsilon} \;\; \boldsymbol{w}]^T_{(1\times(2M+N+2))},$$

where $\boldsymbol{\epsilon}$ and $\boldsymbol{w}$ are respectively defined in (4.22) and (4.29). The objective function Z is defined as

$$Z = \boldsymbol{f}^T \boldsymbol{z},$$

where $\boldsymbol{f}^T = [0 \; 0 \; \mathbf{1}_{2M\times 1} \; \mathbf{0}_{N\times 1}]_{1\times(2M+N+2)}$. The complete linear program is then formulated as:

$$\begin{aligned} \min \; Z &= \boldsymbol{f}^T \boldsymbol{z}, \quad \text{such that} \\ \boldsymbol{D}\boldsymbol{z} &= \boldsymbol{g}, \; \boldsymbol{z} \geq \mathbf{0}, \end{aligned}$$

where

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{A} \mid & \boldsymbol{J} & \mathbf{0}_{M\times N} \\ \boldsymbol{B} \mid & \mathbf{0}_{2M\times N} & \boldsymbol{I}_{N\times N} \end{bmatrix}_{(M+N)\times(2+2M+N)}, \; \boldsymbol{g} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{r} \end{bmatrix}_{(2+2M+N)\times 1}.$$

In the above equations, the matrices $\boldsymbol{A}$, $\boldsymbol{J}$ and $\boldsymbol{B}$ are respectively defined in (4.18), (4.24) and (4.28), and the vectors $\boldsymbol{c}$ and $\boldsymbol{r}$ are defined in (4.19) and (4.28) respectively.

It must be pointed out that in the above linear program, LOS range information is used to define both the objective function and the feasible region, whereas the NLOS range information is used only to define the feasible region. This allows the NLOS range estimates to "assist" in improving the accuracy of location estimates by limiting the size of the feasible region, but does not allow the NLOS bias errors to adversely affect node localization accuracy, since the NLOS range information plays no part in defining the objective function. The efficacy of the proposed method is demonstrated through simulations in section 4.9. The above approach works for any mixture of LOS and NLOS range estimates, provided $m_L \geq 3$, $m_N \geq 0$. In the following section, we discuss some special sub-cases where there are insufficient LOS and NLOS range estimates to apply the approach described above.

4.8 Extensions, Special Cases and Analysis

In the previous section, we formulated the "basic-LP", which utilizes NLOS constraints to generate a rectangular feasible region to limit the potential values for $\hat{\boldsymbol{x}}$. Intuitively, reducing the size of the feasible region can limit large values of the localization error, improving average localization accuracy. In this section, we consider extensions of the basic-LP that utilize additional constraints to further reduce the size of the feasible region. We also consider several special cases where the number of LOS range estimates are not sufficient to apply

the conventional LS solution. We then present an analytical perspective on the impact of the size of the feasible region on the performance of the LP approach.

4.8.1 Piece-wise Linear Feasible Region

Figure 4.10(b) illustrates the use of the tangents at the points of intersection of the circular constraints in reducing the size of the feasible region. Given $m_N > 1$ NLOS constraints, there are a maximum of $L = 2\binom{m_N}{2}$ points of intersection. Consequently, as a tangent to each circle can be drawn at each intersection point, there are at most $2L$ tangents that can be drawn to confine the feasible region. If the points of intersection of the NLOS circular constraints defined in (4.25) are denoted by $\boldsymbol{s}_i = [s_{xi}\ s_{yi}]^T$, $i = 1, 2, \cdots, L$, then a feasible region can be constructed using the convex hull [68] of these points given by

$$F = \left\{ \boldsymbol{x} \,\middle|\, \boldsymbol{x} = \sum_{i=1}^{L} \lambda_i \boldsymbol{s}_i,\ 0 \leq \lambda_i \leq 1,\ \forall i \right\}. \tag{4.30}$$

It must be noted that, as seen in Figure 4.10(b), the above feasible region F is different from the original feasible region discussed in section 4.7.3, and can either be used instead of, or in addition to, the previously defined feasible region. In order to incorporate the above constraints into the original linear program, we can restructure the constraints in (4.30) into the following form:

$$x - \sum_{i=1}^{L} \lambda_i s_{xi} \ = \ 0,\ y - \sum_{i=1}^{L} \lambda_i s_{yi} = 0,\ 0 \leq \lambda_i \leq 1,\ i = 1, 2, \cdots, L.$$

The vector of variables in this case is

$$\boldsymbol{z} = \left[x\ y\ \boldsymbol{l}^T\right]^T_{1\times(L+2)},\ \boldsymbol{l} = [\lambda_1\ \lambda_2\ \cdots\ \lambda_L]^T_{1\times L}.$$

The above constraints can be written in matrix form using:

$$\boldsymbol{H}\boldsymbol{z} = \boldsymbol{0}_{2\times 1},\ \boldsymbol{z} \geq \boldsymbol{0},\ \boldsymbol{l} \leq \boldsymbol{1}_{L\times 1},$$

where

$$\boldsymbol{H} = \begin{bmatrix} 1 & 0 & -s_{x1} & -s_{x2} & \cdots & -s_{xL} \\ 0 & 1 & -s_{y1} & -s_{y2} & \cdots & -s_{yL} \end{bmatrix} = [\boldsymbol{I}_{2\times 2}\ \boldsymbol{S}_{2\times L}],$$

and

$$\boldsymbol{S}_{2\times L} = \begin{bmatrix} -s_{x1} & -s_{x2} & \cdots & -s_{xL} \\ -s_{y1} & -s_{y2} & \cdots & -s_{yL} \end{bmatrix}.$$

4.8.2 Extension to LOS Estimates

In the development of the basic LP, while LOS range estimates are used to define the objective function and the feasible region of the linear program, the NLOS range estimates contribute solely to constraining the feasible region. In the absence of NLOS range estimates ($m_N = 0$, $m_L \geq 3$), it has been verified [27] that the performance of the LP approach in terms of the mean and standard deviation of the localization error is similar to that of the LS estimator. However, given only LOS range estimates, we can artificially introduce a bias to create "artificial" NLOS range estimates which can then be used to confine the feasible region. The underlying assumption here is that a rough estimate of (or at least an upper bound on) the standard deviation of the LOS range estimates is known, such that a large enough bias (sufficiently larger than the standard deviation of the LOS range estimates) can be determined. Thus, even in the absence of "true" NLOS range estimates, we have a method of restricting the feasible region, thereby providing performance gains relative to the LS approach.

If $\boldsymbol{r}_L$ is the given vector of LOS range estimates, introducing a sufficiently large bias results in the "artificial NLOS" vector of range estimates:

$$\tilde{\boldsymbol{r}}_L = \boldsymbol{r}_L + \tilde{\boldsymbol{b}},$$

where $\tilde{\boldsymbol{b}}$ denotes the vector of introduced bias terms. This serves as the vector of NLOS range estimates when $m_N = 0$, and the linear program is formulated exactly as in the case of the basic LP, developed in section 4.7.3, by replacing r_{Nj} with $\tilde{r}_{Lj}$. When $m_N > 0$, $\tilde{\boldsymbol{r}}_L$ is concatenated with $\boldsymbol{r}_N$ resulting in constraints given by (4.28), where $\boldsymbol{B}$, $\boldsymbol{r}$ and $\boldsymbol{w}$ are respectively replaced by:

$$\tilde{\boldsymbol{B}} = \begin{bmatrix} \boldsymbol{B}_1^T & \boldsymbol{B}_1^T & \cdots & \boldsymbol{B}_1^T \end{bmatrix}^T_{(2\times\tilde{N})}, \; \tilde{\boldsymbol{r}} = \begin{bmatrix} \boldsymbol{r}^T & \tilde{\boldsymbol{r}}_L^T \end{bmatrix}^T_{(1\times\tilde{N})}, \tag{4.31}$$

$$\tilde{\boldsymbol{w}} = \begin{bmatrix} \boldsymbol{w}_1^T & \boldsymbol{w}_2^T & \cdots & \boldsymbol{w}_{m_N+m_L}^T \end{bmatrix}^T_{(1\times\tilde{N})}, \tag{4.32}$$

and $\tilde{N} = 4(m_N + m_L)$. In the following section, we present a formulation of the linear program, complete with the extensions discussed above.

4.8.3 Complete Linear Program

Based on the above sections, we can now formulate the complete "extended LP". In this case, the feasible region is determined by (a) both NLOS and "artificial NLOS" constraints, as well as (b) the tangents at intersections of NLOS and "artificial NLOS" circular constraints. The formulation of the extended linear program is as follows: define the vector of variables as

$$\boldsymbol{z} = \begin{bmatrix} x & y & \boldsymbol{\epsilon} & \tilde{\boldsymbol{w}} & \tilde{\boldsymbol{l}} \end{bmatrix}^T_{(1\times(2+2M+\tilde{N}+\tilde{L}))},$$

where $\boldsymbol{\epsilon}$ and $\tilde{\boldsymbol{w}}$ are respectively defined in (4.22) and (4.32). In this case, $\tilde{\boldsymbol{l}} = [\lambda_1\ \lambda_2\ \cdots\ \lambda_{\tilde{L}}]^T_{1\times\tilde{L}}$, $\tilde{L} = 2\binom{m_N+m_L}{2}$. The objective function Z is defined as

$$Z = \boldsymbol{f}^T\boldsymbol{z},$$

where $\boldsymbol{f}^T = \left[0\ 0\ \mathbf{1}_{2M\times 1}\ \mathbf{0}_{\tilde{N}\times 1}\ \mathbf{0}_{\tilde{L}\times 1}\right]_{1\times\left(2+2M+\tilde{N}+\tilde{L}\right)}$. The linear program is then formulated as:

$$\begin{aligned} \min\ Z &= \boldsymbol{f}^T\boldsymbol{z}, \quad \text{such that} \\ \boldsymbol{D}\boldsymbol{z} &= \boldsymbol{g},\ \boldsymbol{z} \geq \mathbf{0},\ \tilde{\boldsymbol{l}} \leq \mathbf{1}_{\tilde{L}\times 1}, \end{aligned}$$

where

$$\boldsymbol{D} = \begin{bmatrix} \boldsymbol{A}_{M\times 2} \mid & \boldsymbol{J}_{M\times 2M} & \mathbf{0}_{M\times\tilde{N}} & \mathbf{0}_{M\times\tilde{L}} \\ \boldsymbol{B}_{\tilde{N}\times 2} \mid & \mathbf{0}_{\tilde{N}\times 2M} & \boldsymbol{I}_{\tilde{N}\times\tilde{N}} & \mathbf{0}_{\tilde{N}\times\tilde{L}} \\ \boldsymbol{I}_{2\times 2} \mid & \mathbf{0}_{2\times 2M} & \mathbf{0}_{2\times\tilde{N}} & \boldsymbol{S}_{2\times\tilde{L}} \end{bmatrix}_{(M+\tilde{N}+2)\times(2+2M+\tilde{N}+\tilde{L})}, \boldsymbol{g} = \begin{bmatrix} \boldsymbol{c} \\ \boldsymbol{r} \\ \mathbf{0}_{\tilde{L}\times 1} \end{bmatrix}_{\left(2+2M+\tilde{N}+\tilde{L}\right)\times 1}.$$

4.8.4 Special Cases

The extended linear program described above can be used to accurately estimate a node's location for a large number of cases with a mixture of LOS and NLOS range estimates, when $m_L \geq 3$, $m_N \geq 0$. In order to generalize the scheme to handle other *degenerate* cases, we consider the exceptions below:

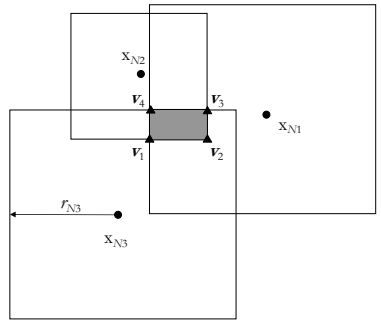
Case I: $m_L = 0$**,** $m_N > 0$

If no LOS range estimates are available, then the LP approach in the form discussed above will not be applicable since the NLOS range estimates are not used to define an objective function. An example of this situation with $m_N = 3$ is shown in Figure 4.11(a). In this case (if $m_N \geq 3$), we could either use the LS estimator without the mitigation of bias errors, or simply use the centroid of the feasible region as a location estimate:

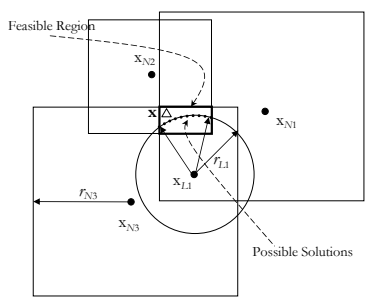
$$\hat{\boldsymbol{x}} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{v}_i,$$

where $\boldsymbol{v}_i$, $i = 1, 2, \cdots, n$, are the vertices of the feasible region. In the case of the basic LP, the feasible region is rectangular ($n = 4$), and it is straightforward to show that:

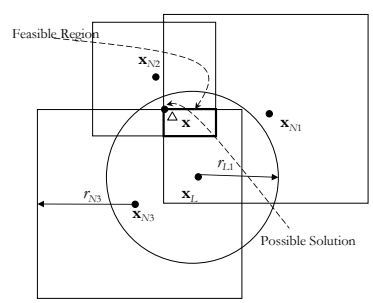
$$\hat{\boldsymbol{x}} = \frac{1}{2}\begin{bmatrix} \min_i\{x_{Ni} + r_{Ni}\} + \max_i\{x_{Ni} - r_{Ni}\} \\ \min_i\{y_{Ni} + r_{Ni}\} + \max_i\{y_{Ni} - r_{Ni}\} \end{bmatrix}.$$



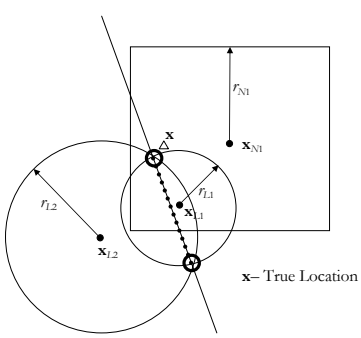
(a) Case I: $m_L = 0$ and $m_N = 3$. The vertices of the feasible region are denoted by $\boldsymbol{v}_1$, $\boldsymbol{v}_2$, $\boldsymbol{v}_3$ and $\boldsymbol{v}_4$.



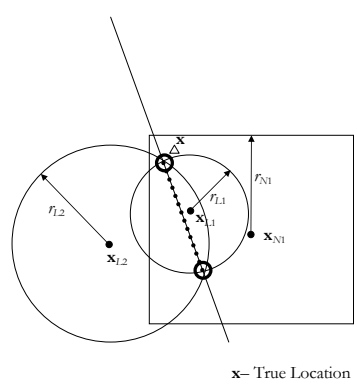
(b) Case II: $m_L = 1$, $m_N \geq 2$. Sub-case 1: the LOS constraint cuts through the feasible region generated using the NLOS constraints.



(c) Case II: $m_L = 1$, $m_N \geq 2$. Sub-case 1: the LOS constraint does not pass through the feasible region generated using the NLOS constraints.



(d) Case III: $m_L = 2$, $m_N \geq 1$, Sub-case 2: The feasible region contains one of two intersection points of the circle.



(e) Case III: $m_L = 2$, $m_N \geq 1$, Sub-case 3: The feasible region contains both intersection points of the circle.

Case II: $m_L = 1$**,** $m_N \geq 2$

This situation is illustrated in Figures 4.11(b) and 4.11(c), which represent two sub-cases. Since $m_L = 1$, we have a single LOS equality constraint and the linearization performed in (4.15) is not possible. In the first sub-case (Figure 4.11(b)), the circle formed using the LOS range estimate r_{L1} passes through the feasible region formed by the NLOS constraints (and with LOS constraints in the extended LP). In such a case, the center of the arc of the circle within the feasible region can be selected as the solution. In the second sub-case, illustrated in Figure 4.11(c), the circular constraint formed using the LOS range estimate does not pass through the feasible region. In such a case, we can pick the vertex of the feasible region that is closest to the circle as a potential solution, i.e.,

$$\hat{\boldsymbol{x}} = \arg\min_{\boldsymbol{v}_i} \left\{ \mid \|\boldsymbol{v}_i - \boldsymbol{x}_{L1}\| - r_{L1} \mid \right\}.$$

Case III: $m_L = 2$**,** $m_N \geq 1$

In this case, since $m_L = 2$, we have two circular constraints, and the linearization of LOS estimates is not particularly useful since a single linear constraint is generated using the difference. Therefore, instead of two potential solutions corresponding to the intersections of the two circles, we have an infinite number of solutions. It is easier to compute the two intersections of the circles, and if there are additional NLOS constraints, three sub-cases arise: (1) Neither of the intersection points lies inside the feasible region formed by the NLOS constraints, (2) only one of the intersection points lies inside the feasible region (Figure 4.11(d)), and (3) both intersection points lie inside the feasible region formed by the NLOS constraints (Figure 4.11(e)). When neither intersection point lies inside the feasible region formed by the NLOS constraints, we can pick the intersection that lies closest to the centroid of the feasible region. In the second sub-case, we can simply pick the intersection point that lies within the feasible region. In the last sub-case, in order to eliminate one of the solutions, consider the following approach: since the NLOS range estimates are larger than the true ranges, decrement the NLOS range estimates by small amounts until only one of the intersection points remains inside the feasible region. The intersection point remaining in the feasible region is likely to be closer to the true location than the intersection point that lies outside the feasible region, and can be selected as the location estimate.

4.8.5 Other Extensions

Some furthers extensions of the linear program are possible. For example, similar to the heuristic weighted LS method discussed in Section 3.5.1, we can weight the terms in the objective function of the linear program, based on a factor proportional to the LOS range estimates. The rationale for this weighting is that the variance of the range estimate increases

with distance. The objective function of the linear program with weighting can be written as:

$$Z = \sum_{i} \sum_{j,\ j>i} \gamma_{ij} \left(e_{ij}^{+} + e_{ij}^{-} \right), \tag{4.33}$$

where γ_{ij} is the weight corresponding to the linear equation extracted through the difference between the ith and jth LOS range estimates. A possible set of weights is $\gamma_{ij} = r_{Li} r_{Lj}$. Note that methods is not strictly analogous to the heuristic weighting method of Section 3.5.1, as it includes differences of *all* pairs of LOS range estimates.

4.8.6 Impact of the number of NLOS range estimates

The efficacy of the proposed approach depends on the fact that the NLOS bias errors are much larger than the range measurement noise, so that the true node location $\boldsymbol{x}$ lies inside the feasible region generated using the NLOS range estimates. Given that the true node location lies inside the feasible region, it is evident that reducing the size of feasible region improves localization accuracy by limiting the localization error. As the number of NLOS range estimates m_N increases or as $B_{\max}$ decreases, two trends which have opposing effects on the localization error are expected: (a) the probability p that the true solution lies outside the feasible region formed by NLOS estimates increases, and (b) the area A of the feasible region decreases. If the true solution lies inside the feasible region, then reducing the area of the feasible region also reduces the mean and variance of the localization error. It is important to point out that even if the true location does not lie inside the feasible region, this does not automatically imply a large localization error (especially when the feasible region is small), although the two events are correlated. This suggests that the area of feasible region A may play a more significant role in determining average localization error than the probability p.

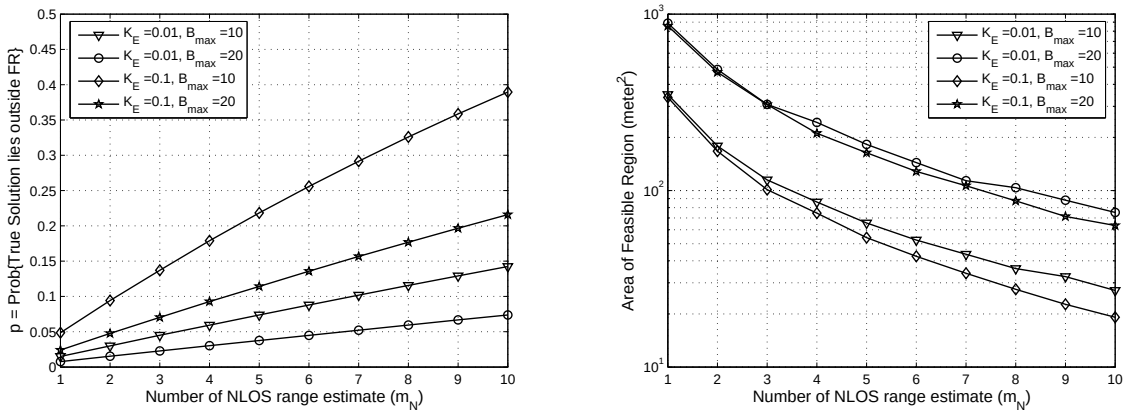
The probability p that $\boldsymbol{x}$ lies outside the linear feasible region formed using NLOS range estimates can be upper-bounded by the probability that the true distance $\|\boldsymbol{x} - \boldsymbol{x}_{Nj}\|$ is larger than the range estimate r_{Nj} for at least one value of j, $j \in \{1, 2, \cdots, m_N\}$. In such a case, the region of overlap between the NLOS constraints does not contain the true location $\boldsymbol{x}$. From the model for NLOS range estimates given in (4.12), p can be bounded by:

$$p \leq 1 - \prod_{j=1}^{m_N} q_j, \tag{4.34}$$

where q_j is given by:

$$q_j = \Pr\left\{ r_{Nj} < R_{Nj} \right\} = \Pr\left\{ b_{Nj} + n_{Nj} < 0 \right\},\ j = 1, 2, \cdots, m_N.$$

It must be emphasized that (4.34) is an upper-bound for p, since the right-hand-side of (4.34) represents the probability that the true solution lies outside the original feasible region formed



(a) (b)

Figure 4.12: **(a)** The upper bound on the probability p that the true solution lies outside the feasible region generated using m_N NLOS range estimates for different values of K_E and $B_{\max}$. **(b)** The area A of the feasible region as m_N increases for different values of K_E and $B_{\max}$.

using circular constraints, which is contained within the linearized feasible region. Assuming that $n_{Nj} \sim \mathcal{N}(0, \sigma_{Nj}^2)$ and $b_{Nj} \sim \mathcal{U}(0, B_{\max})$, we can show that (see Appendix 4C):

$$q_j = 1 - \frac{1}{2}\text{erfc}\left(\frac{z_j}{\sqrt{2}}\right) - \frac{1}{\sqrt{2\pi}}\left(\frac{1 - e^{-\frac{z_j^2}{2}}}{z_j}\right), \; z_j = \frac{B_{\max}}{\sigma_{Nj}}. \tag{4.35}$$

Figure 4.12(a) shows the variation of the probability p, computed using (4.34) and (4.35), as m_N increases, with the simulation parameters used in Figures 4.8(a) and 4.8(b). Figure 4.12(b) shows the decrease in the area of the feasible region as the number of available NLOS range estimates increases. We see that as $B_{\max}$ increases, the area of the feasible region is larger, implying reduced gains due to NLOS constraints on the feasible region, but the probability $(1-p)$ of the feasible region containing the true location increases. Further, we do observe that when K_E (in (4.11)) is small, the probability p is small and therefore, we are likely to see a considerable decrease in the mean localization error as the number of NLOS range estimates increases. Further, we expect that the extent of the improvement obtained by restricting the feasible region with additional NLOS range estimates diminishes as the number of LOS range estimates increases.

As we will observe from the simulation results presented in the following section, decreasing $B_{\max}$ or increasing m_N reduces the area of the feasible region A and can significantly decrease average localization error, despite the resulting increase in p.

4.9 Simulation Results

In this section, we present simulation results that demonstrate that the proposed approach not only mitigates the effect of NLOS bias errors but utilizes the NLOS range information to improve node localization accuracy. In the following discussion of simulation results, the anchor nodes are randomly (uniformly) distributed over a $W \times W$ area, where $W = 10$ meters. The unknown node location of interest is $\boldsymbol{x} = [5\ 5]^T$ (meters). The values of β_L and β_N are assumed to be 2 and 2.5 respectively. We compare the performance of four location-estimation approaches in terms of the mean and standard deviation of the localization error Ω: (i) the basic LP approach ("LP-Basic") discussed in section 4.7.3, (ii) the extended LP approach ("LP-Extended") discussed in section 4.8.3, (iii) the LS estimator, utilizing only LOS range estimates, while discarding the NLOS range estimates ("LS-(Pure-LOS)"), and (iv) the LS estimator, utilizing both LOS and NLOS range estimates, without the mitigation of NLOS bias errors ("LS-(LOS+NLOS)"). For the extended LP approach, the standard deviation of LOS range estimates is assumed to be known, and bias errors equal to 3 times the standard deviation are added to generate the artificial NLOS range estimates.

For these four methods, the values of Ω are computed for a large number ($N_{iter} = 20000$) of realizations of the measurement noise and bias errors for 50 sets of randomly generated anchor locations. The mean μ_Ω and the standard deviation σ_Ω of the localization error are shown in Figure 4.13(a), for different values of the proportionality constant K_E defined in (4.11). In this simulation, $m_L = 3$, $m_N = 3$ and $B_{\max} = 8$ meters. As expected, for all location estimators, node localization accuracy degrades as the variance of the range estimates increases (i.e., K_E increases). The proposed LP approaches outperform the LS-based schemes in terms of both the mean and standard deviation of the localization error, and therefore, on the average, generate more accurate node location estimates. In general, it is observed that for the mentioned estimation procedures, μ_Ω and σ_Ω follow identical trends in terms of their variation with K_E, even for larger values of β_N:

$$\begin{aligned} \mu_{\Omega,LP-Extended} &< \mu_{\Omega,LP-Basic} < \mu_{\Omega,LS-Pure\ LOS} < \mu_{\Omega,LS-(LOS+NLOS)}, \\ \sigma_{\Omega,LP-Extended} &< \sigma_{\Omega,LP-Basic} < \sigma_{\Omega,LS-Pure\ LOS} < \sigma_{\Omega,LS-(LOS+NLOS)}. \end{aligned} \tag{4.36}$$

The variation of the mean localization error μ_Ω with K_E, while respectively increasing the maximum bias $B_{\max}$, and the number of NLOS range estimates m_N, is shown in Figures 4.13(b) and 4.13(c). We observe that (a) the LP approaches outperform both the LS approaches and are less sensitive to an increase in $B_{\max}$, and (b) the performance of the LP approaches improves as m_N increases. The former effect is due to the fact that the NLOS ranges do not contribute to the objective function of the linear program, while the latter is because additional NLOS range estimates reduce the size of the feasible region, thereby limiting large values of the localization error. The LP approaches once again outperform the LS estimator that utilizes only LOS range estimates.

The results in Figures 4.13(b) and 4.13(c) are in agreement with the analysis presented

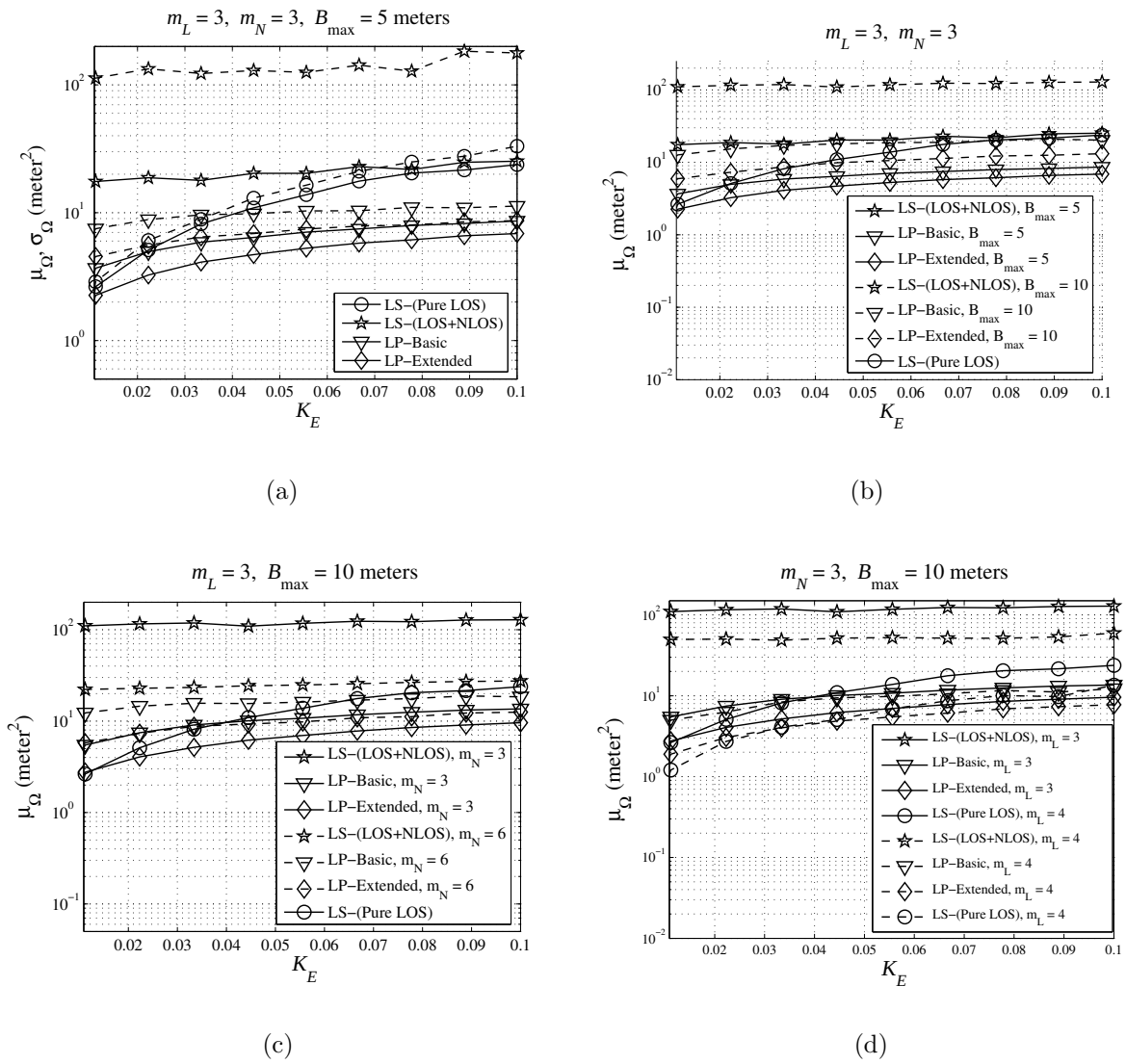


Figure 4.13: **(a)** The mean μ_Ω and the standard deviation σ_Ω of the localization error Ω are plotted versus K_E. Here, $m_L = 3$, $m_N = 3$ and $B_{\max} = 5$ meters. **(b)** Mean μ_Ω of the Localization Error Ω, $m_L = 3$, $m_N = 3$. The maximum bias $B_{\max}$ is increased from 5 meters to 10 meters. **(c)** Mean μ_Ω of the Localization Error Ω, $m_L = 3$, $B_{\max} = 10$ meters. The number of NLOS range estimates is varied from $m_N = 3$ to $m_N = 6$. **(d)** Mean μ_Ω of the Localization Error Ω, $m_N = 4$, $B_{\max} = 10$ meters, as the number of LOS range estimates is increased from $m_L = 3$ to $m_L = 4$.

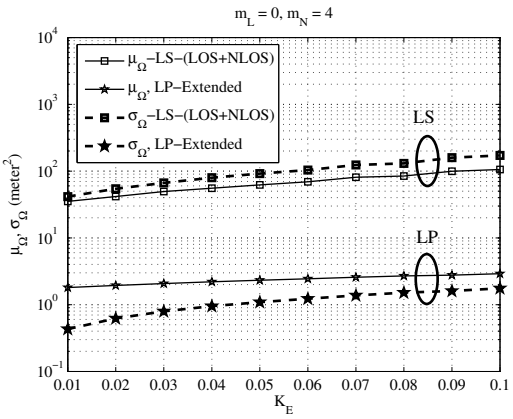
in Section 4.8.6, where we argued that either decreasing $B_{\max}$ or increasing m_N results in the reduction of the size of the feasible region and the average localization error, despite the increase in the probability that the true solution lies outside the feasible region. We also note that the extended-LP approach outperforms the basic-LP approach, due to the availability of a larger number of constraints on the feasible region, at the cost of higher complexity.

From Figure 4.13(d), we see that the improvement in performance of the LP-based approaches over the LS-(Pure LOS) approach decreases as the number of LOS range estimates increases. For instance, when $K_E = 0.1$, the percentage improvement of the extended LP method over the LS-(Pure LOS) estimator decreases from 78% when $m_L = 3$ to 65% when $m_L = 4$. Therefore, when the number of LOS range estimates is large ($m_L \gg 3$), discarding the NLOS range estimates is likely to have negligible impact on localization accuracy, and the LS-(Pure LOS) estimator can be used to generate accurate location estimates.

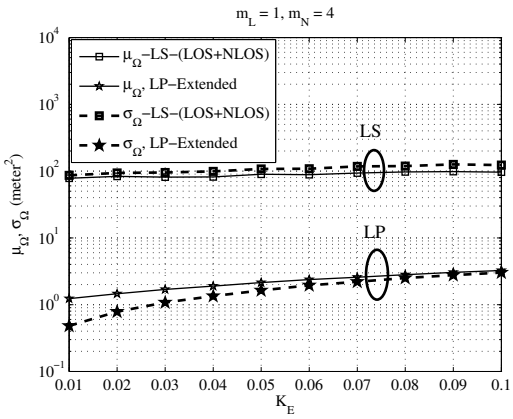
Figures 4.14(a)-4.14(c) compare the performance of the extended LP approach with the LS-(LOS+NLOS) approach in scenarios where the LS-(Pure LOS) approach cannot be applied due to an insufficient number of LOS range estimates ($m_L < 3$), i.e., the special cases discussed in Section 4.8.4. In each of these cases, we see that the LP approach substantially outperforms the LS approach in terms of μ_Ω and σ_Ω.

The simulation results presented thus far assumed that the NLOS bias errors were uniformly distributed between 0 and $B_{\max}$. Figure 4.14(d) compares the performance of the approaches when the NLOS bias errors are exponentially distributed [44], which has been observed in UWB NLOS measurements [1], [43]. For the sake of comparison, the mean NLOS bias for both uniform and exponential distributions was set at $B_{\max}/2$, and the mean localization error was computed for 100 random anchor locations, each with a large number of range measurement noise realizations. We see that the even when the bias errors are exponentially distributed, the LP-based approaches provide significant reduction in the average localization error as compared to the LS-based approaches. In fact, the LP-based approaches perform approximately 20% better when the bias errors are exponentially distributed than when they are uniformly distributed. With exponentially distributed bias errors, the probability of obtaining smaller bias values is higher than with uniformly distributed bias errors; smaller bias values imply a smaller feasible region which, in turn, results in lower average localization error. Finally, this result suggests the possibility that the LP-based approaches can provide gains in localization accuracy relative to LS estimators in general, when empirical statistical distributions of NLOS bias errors are applied.

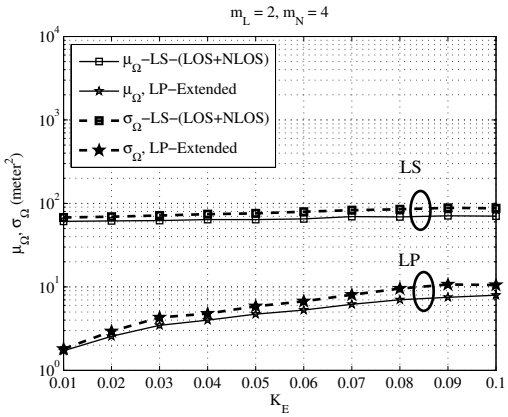
Figures 4.15(a) and 4.15(b) compare the performance of the LS - (Pure LOS) and LP - Extended estimators for different values of m_L and m_N. From Figure 4.15(a), we see that for $m_L > 6$, the LS-(Pure LOS) estimator outperforms the LP-Extended estimator, and thus, beyond this value of m_L, discarding NLOS range estimates is advisable. As m_L increases, the probability p increases leading to performance degradation relative to the LS-(Pure LOS) estimator. As similar effect is seen with an increasing m_N in Figure 4.15(b). We see that for each value of m_L there is an optimal value of m_N beyond which performance degrades as



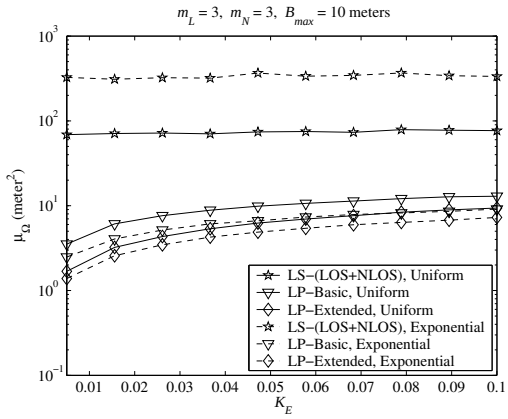
(a) $m_L = 0$, $m_N = 4$



(b) $m_L = 1$, $m_N = 4$



(c) $m_L = 2$, $m_N = 4$



(d) $m_L = 3$, $m_N = 4$, Exponentially and uniformly distributed NLOS bias errors

Figure 4.14: Some sub-cases where discarding the NLOS estimates is not possible when the LS estimator is used, as $m_L < 3$. **(a)** Case I: $m_L = 0$, $m_N = 4$, **(b)** Case II: $m_L = 1$, $m_N = 4$, **(c)** Case III: $m_L = 2$, $m_N = 4$. **(d)** A comparison of the mean localization error obtained using different approaches when the NLOS bias errors are (i) uniformly distributed and (ii) exponentially distributed. The mean bias error for both distributions is normalized to $B_{\max}/2$, where $B_{\max} = 10$ meters.

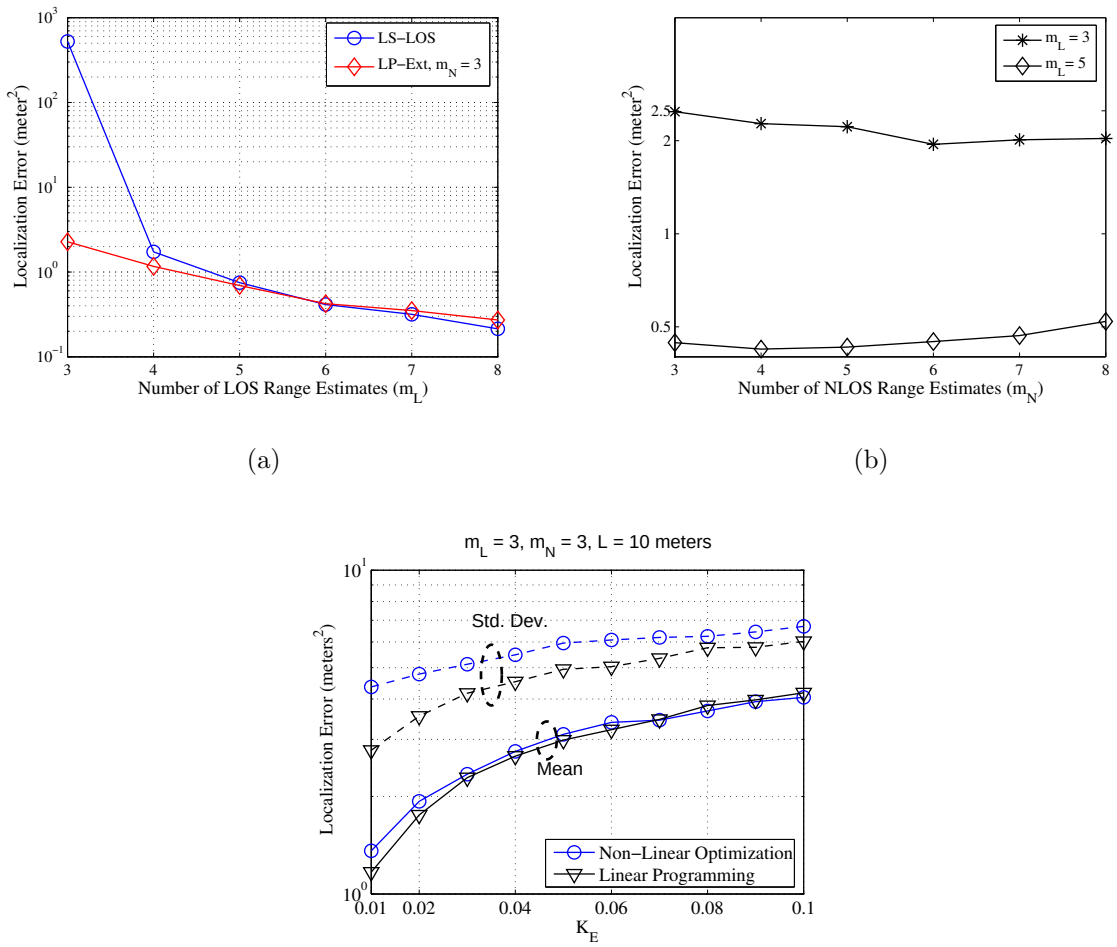


Figure 4.15: (a) Comparison of the LS-(Pure LOS) and the LP-Extended estimators for different values of m_L with $m_N = 3$, (b) Performance of the LP-Extended estimator for different values of m_N with $m_L = \{3, 5\}$, (c) Comparison of the LP-Extended method with the non-linear constrained optimization approach when $m_L = 3$, $m_N = 3$. In each of the above simulations, $L = 10$ meters, $\beta = 2$, and $x = [5\ 5]^T$.

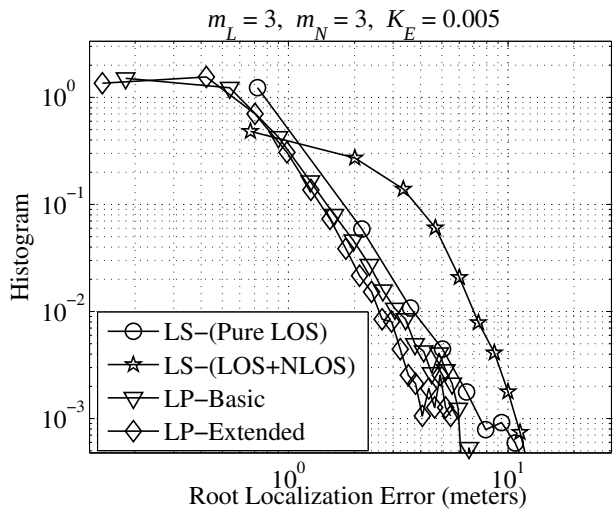


Figure 4.16: Histogram of the root localization error (meters) using different methods when $K_E = 0.005$, $m_L = 3$, $m_N = 3$, $B_{\max} = 10$ meters.

m_N increases. This optimal value of m_N decreases as m_L increases, e.g., $m_{N,\ opt} = 6$ when $m_L = 3$, and $m_{N,\ opt} = 4$ when $m_L = 5$. This effect is again attributable to the probability p. As either m_N or m_L increase, the total number of effective NLOS constraints in the LP methods increase. This reduces the area of the feasible region A but increases p, and beyond a certain point the increase in p dictates performance. Finally, Figure 4.15(c) compares the performance of the LP-Extended method with non-linear constrained optimization (implemented using the fmincon routine in Matlab) for $m_L = 3$, $m_N = 3$. We see that the LP method and the non-linear estimator are comparable in terms of the mean localization error, whereas the LP method outperforms the non-linear estimator in terms of the standard deviation of the localization error.

One of the main drawbacks of using the mean and standard deviation of the localization error as indicators of the localization accuracy is that a single realization of noise or node locations that results in a large error can considerably increase the values of μ_Ω and σ_Ω. This effect is particularly observed in the LS-(Pure LOS) scenario, where a single occurrence of large GDOP can result in large values of μ_Ω and σ_Ω. As the number of realizations of the node locations considered increases, the probability of the occurrence of a scenario with poor anchor geometry increases, and consequently, μ_Ω and σ_Ω also increase. Figure 4.16 compares the histograms of the root localization error obtained using different location estimators. We see that the LP-based methods display a higher probability of small localization errors than the LS-based methods. We also notice that the LS-(Pure LOS) displays a significant probability of large localization error which can adversely impact the *average* localization accuracy. However, based on the histogram, the trends in (4.36) are confirmed.

The trends observed in the simulation results are summarized below:

1. In terms of the mean and standard deviation of the localization error, the LP-based estimators outperform the LS-based estimators for small (practical) values of m_L, and when $m_N \geq 1$. The average localization accuracy of the LS-(Pure LOS) estimator is dominated by the effect of high GDOP for small values of m_L.

2. When m_L is large, the LS-(Pure LOS) estimator outperforms the LP-Extended estimator. Thus, when the number of LOS range estimates is large, discarding NLOS range estimates may be advisable.

3. The LP approaches are less sensitive to an increase in the bias of NLOS range estimates than the LS approaches.

4. For each value of m_L, there is an optimal value of m_N beyond which performance of the LP-based estimators degrades as m_N increases. This optimal value of m_N decreases as m_L increases.

5. The LP-based approaches perform better when the bias errors are exponentially distributed than when they are uniformly distributed.

4.10 Simulation Results: Application to Mobile Ad hoc Location-Aware Networks

In the previous section, we investigated the average localization accuracy obtained in the case of a stationary unlocalized node. In this section, we evaluate the efficacy of the proposed NLOS mitigation scheme in a mobile ad hoc network, where localization accuracy is a function of time. It must be pointed out that in mobile scenarios, tracking algorithms can be applied in order to mitigate the impact of NLOS bias errors [69, 70, 71]. However, the efficacy of tracking algorithms may be limited in indoor networks due to the fact that in the case of ad hoc deployment of nodes in a typical indoor environment, the number of NLOS range estimates available can far exceed the number of LOS range estimates *throughout* the duration of the mobile's trajectory. As a result, the tracking algorithm is provided with biased range estimates at every instant of time, and as a consequence, its ability to ameliorate present estimates based on past location and range estimates is reduced. However, tracking algorithms can be used to exploit the spatial correlation between location and range estimates within the mobile's trajectory, and can thus operate in conjunction with NLOS identification and mitigation schemes to improve localization accuracy.

In this simulation framework, the network is assumed to comprise N_A anchors distributed over an area (or volume) of interest, and whose locations are known *a priori*. A single mobile node, whose time-varying location is denoted by the vector $\boldsymbol{x}(t)$, is assumed to travel through this region at a constant speed of v meters per second. The location of the mobile node at

time $t = (n+1)T_s$ is generated using the following linear model:

$$\boldsymbol{x}_M((n+1)T_s) = \boldsymbol{x}_M(nT_s) + v\boldsymbol{e}(nT_s), \quad \boldsymbol{x}_M(0) = \boldsymbol{x}_{M0}, \tag{4.37}$$

where $\boldsymbol{x}_{M0}$ is the location of the mobile at $t = 0$, and $\boldsymbol{e}(t)$ is the time-varying unit vector in the direction of the mobile's instantaneous motion.

The mobile is assumed to gather range estimates from nearby anchors and estimate its location every T_s seconds. At time $t = nT_s$, successful "ranging" is assumed to occur between the mobile node and the ith anchor if the effective signal-to-noise-ratio (SNR) $\xi_i(nT_s)$ is larger than a threshold ξ_T. The effective SNR $\xi_i(nT_s)$ on the link between the ith anchor and mobile node is modeled by

$$\xi_i(nT_s) = K_P P_T R_i^{-\beta}(nT_s),$$

where P_T denotes the (constant) transmit power, $R_i(nT_s) = \|\boldsymbol{x}(nT_s) - \boldsymbol{x}_i\|$ is the distance between the mobile node and the ith anchor and K_P is a constant that subsumes the effects of other physical layer parameters. Depending on the locations of the mobile node, anchors and obstructions, at any instant of time, the mobile node may successfully obtain range estimates from both LOS and NLOS anchors. The path loss exponents for the LOS and NLOS cases are assumed to be β_L and β_N respectively, $\beta_L < \beta_N$. Therefore, at a given instant of time, the mobile node receives range estimates from LOS and NLOS anchors that are respectively within distances

$$R_{\max,L} = \left(\frac{K_P P_T}{\xi_T}\right)^{\frac{1}{\beta_L}}, R_{\max,N} = \left(\frac{K_P P_T}{\xi_T}\right)^{\frac{1}{\beta_N}},$$

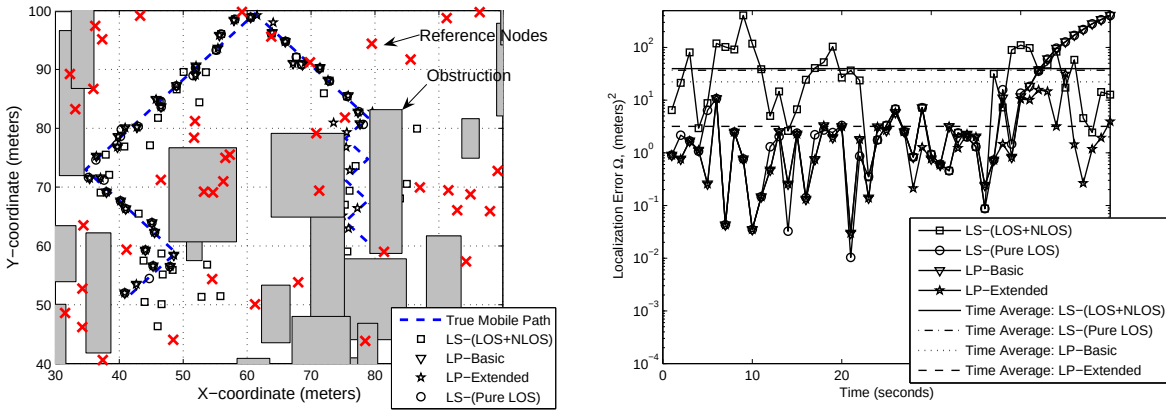
from the mobile's current location. We assume that the nature of the propagation channel (LOS or NLOS) is identified for each received range estimate and, for simplicity, that the range estimates are uncorrelated at different time instants. At time t, let the set of all anchor locations $\mathcal{A}$ be partitioned into the set of locations of all anchors which have LOS to the mobile, $\mathcal{A}_{\mathcal{L}t}$, and the set of locations of all anchors which do not have LOS to the mobile, $\mathcal{A}_{\mathcal{N}t}$, such that $\mathcal{A}_{\mathcal{L}t} \bigcup \mathcal{A}_{\mathcal{N}t} = \mathcal{A}$, $\mathcal{A}_{\mathcal{L}t} \bigcap \mathcal{A}_{\mathcal{N}t} = \phi$. Then, the subsets of LOS and NLOS anchors which provide range estimates at time t are respectively defined by

$$\mathcal{L}_t = \{\boldsymbol{x}_A \,|\boldsymbol{x}_A \in \mathcal{A}_{\mathcal{L}t},\ \|\boldsymbol{x}(t) - \boldsymbol{x}_A\| \leq R_{\max,L}\}, \mathcal{N}_t = \{\boldsymbol{x}_A \,|\boldsymbol{x}_A \in \mathcal{A}_{\mathcal{N}t},\ \|\boldsymbol{x}(t) - \boldsymbol{x}_A\| \leq R_{\max,N}\}.$$

At $t = nT_s$, the location estimate $\hat{\boldsymbol{x}}(t)$ is computed using LOS and NLOS range estimates from $\mathcal{L}_t$ and $\mathcal{N}_t$, governed by (4.10) and (4.12) respectively. The localization error as a function of time is computed using:

$$\Omega(t) = \|\boldsymbol{x}(t) - \hat{\boldsymbol{x}}(t)\|^2,$$

using the different estimation approaches for $\hat{\boldsymbol{x}}(t)$.



(a) (b)

Figure 4.17: Simulation of a 2D location-aware network in a NLOS environment: the values of the simulation parameters used are: $P_T = 1$ mW, $K_P = 1 \times 10^5$, $\xi_T = 15$ dB, $\beta_L = 2$, $\beta_N = 3$, $K_E = 0.1$ and $B_{\max} = 10$ meters. The total number of anchors is $N_A = 80$, and the mobile node moves at a speed of $v = 2.5$ meters/second through a $W \times W$ meter^2 area, where $W = 100$ meters. **(a)** The true mobile trajectory and the mobile location estimates obtained using the LS and the LP-based location estimators. **(b)** A comparison of the different location estimators in terms of the localization error Ω versus time. We observe the trends suggested by equation (4.36).

4.10.1 2-D Application: Indoor Tracking

Consider a $W \times W$ meter^2 area of interest with $W = 100$ meters, containing $N_A = 80$ randomly (uniformly) distributed anchors. Several obstructions are randomly dispersed over the area of interest. At time t, if the path between the mobile's true location $\boldsymbol{x}(t)$ and a given anchor contains any portion of the obstructions, then the range estimates from that anchor are assumed to be given by (4.12); otherwise, the range estimates are given by (4.10). The density of nodes in this simulation was selected such that the total number of range estimates $(m_L + m_N)$ received at each instant of time lies between 3 and 8. The mobile is assumed to move at a speed $v = 2.5$ meters per second in a direction $\boldsymbol{e}(0) = \frac{1}{\sqrt{2}}[1\ 1]^T$ at $t = 0$ and subsequently reflects off the boundary of any obstacle encountered.

Figure 4.17(a) compares the true location of a mobile node with the location estimates computed using the LS and LP approaches. The mobile node estimates its location based on range information from anchors every $T_s = 1$ second between $t = 0$ and $t = 50$ seconds. We see that the LP-based NLOS mitigation schemes outperform the two LS-based schemes. This is more evident in Figure 4.17(b) which compares the localization error achieved by the estimation approaches versus time. We see that on the average, the LP-based NLOS mitigation schemes achieve much higher localization accuracy than the LS based location estimators.

4.10.2 3-D Application: Indoor tracking with ground sensors

Consider a building with 3 floors, each floor having dimensions $W \times W \times H$, $W = 100$ meters, $H = 10$ meters, as shown in Figure 4.18(a). Let a total of $N_A = 108$ anchors be deployed uniformly over the 3 floors on the ground, i.e., $\frac{N_A}{3} = 36$ anchors on each floor. The coordinates of the anchors in meters are generated using

$$\boldsymbol{x}_A = \left\{ [L \cdot u_1 \quad L \cdot u_2 \quad (g + F \cdot H)]^T \right\}$$

where u_1 and u_2 are uniformly distributed random variables over $[0, 1]$, g is a zero-mean Gaussian random variable with unit variance and $F = \{0, 1, 2\}$ is the index of the floor. A single mobile node is assumed to be present on the floor $F = 1$, moving linearly with a speed $v = 1.5$ meters per second from an initial location $\boldsymbol{x}_{M0} = [5\ 5\ H]^T$ (meters) at $t = 0$. The motion of the mobile is defined using (4.37), where

$$\boldsymbol{e}(nT_s) = \frac{1}{2}[\cos(0.2\pi nT_s)\ \sin(0.2\pi nT_s)\ 1]^T;$$

It is assumed that all anchors on floor $F = 1$ have LOS links to the mobile, and provide unbiased range estimates. Further, anchors from floors $F = \{0, 2\}$ are NLOS with respect to the mobile node, and therefore provide biased range estimates.

In this scenario, all the LOS anchors are approximately located on the plane $z = H$. As a consequence, it is expected that the LS-(Pure LOS) approach will have poor location resolution along the z-direction relative to the x and y directions. In such a case, it is intuitive that the NLOS mitigation approaches would play a key role in determining localization accuracy along the z-direction, as the NLOS anchors are located around different planes $z = 0$ and $z = 2H$, and can provide improved resolution in the z-direction. This intuition is confirmed in Figures 4.18(b) and 4.18(c), where we see that any scheme that utilizes NLOS information provides better location resolution in the z-direction. From Figure 4.18(c), we see that the LS-(Pure LOS) method has very poor localization accuracy in the z-coordinate of the mobile's location. Due to the presence of a large number of LOS anchors along the xy-plane at $z = H$, the discussed approaches perform similarly in terms of localization accuracy in the x and y coordinates of the mobile as seen in Figure 4.18(b).

Figure 4.18(d) compares the different schemes in terms of the total localization error $\Omega(t)$ as the mobile moves through the volume of interest at a speed v between $t = 0$ and $t = 25$ seconds. The mobile attempts to compute its location every $T_s = 0.5$ seconds. Once again, we see that the LP-based localization schemes outperform the LS-based schemes in terms of the time-averaged localization error by an order of magnitude. Note that this is a particularly interesting scenario since 3D-localization is crucial to several applications, but is known to be difficult in indoor scenarios. This example highlights the potential benefit of the proposed approach in practical scenarios.

4.11 Conclusions from Part II

In Part II, we described a novel linear-programming approach to the problem of localization in NLOS environments. The main motivation for the development of this method was that in typical indoor PoLoNets, it is likely that we would be required to compute a node's location using a mixture of LOS and NLOS range estimates. Using the LOS range estimates to define the objective function, and the NLOS range estimates to restrict the feasible region for the linear program, we showed that NLOS range information can be used to improve node localization accuracy without incurring performance degradation due to NLOS bias errors. This approach was observed to perform as well as the LS estimator when only LOS range estimates are provided, and substantially better than the LS estimator when a mixture of LOS and NLOS range estimates is provided, particularly when the number of LOS range estimates is small. Relative to LS estimation, it was found that the LP method is less sensitive to increase in NLOS bias errors, and that increasing the number of NLOS range estimates improves node localization accuracy. Further, the proposed approach can be applied to the general (and practically significant) three-dimensional location-estimation problem, and to a number of degenerate cases with insufficient LOS range estimates.

The joint efficacy of the proposed NLOS identification and mitigation methods will be demonstrated in a location-tracking experiment with indoor UWB measurements in Chapter 6.

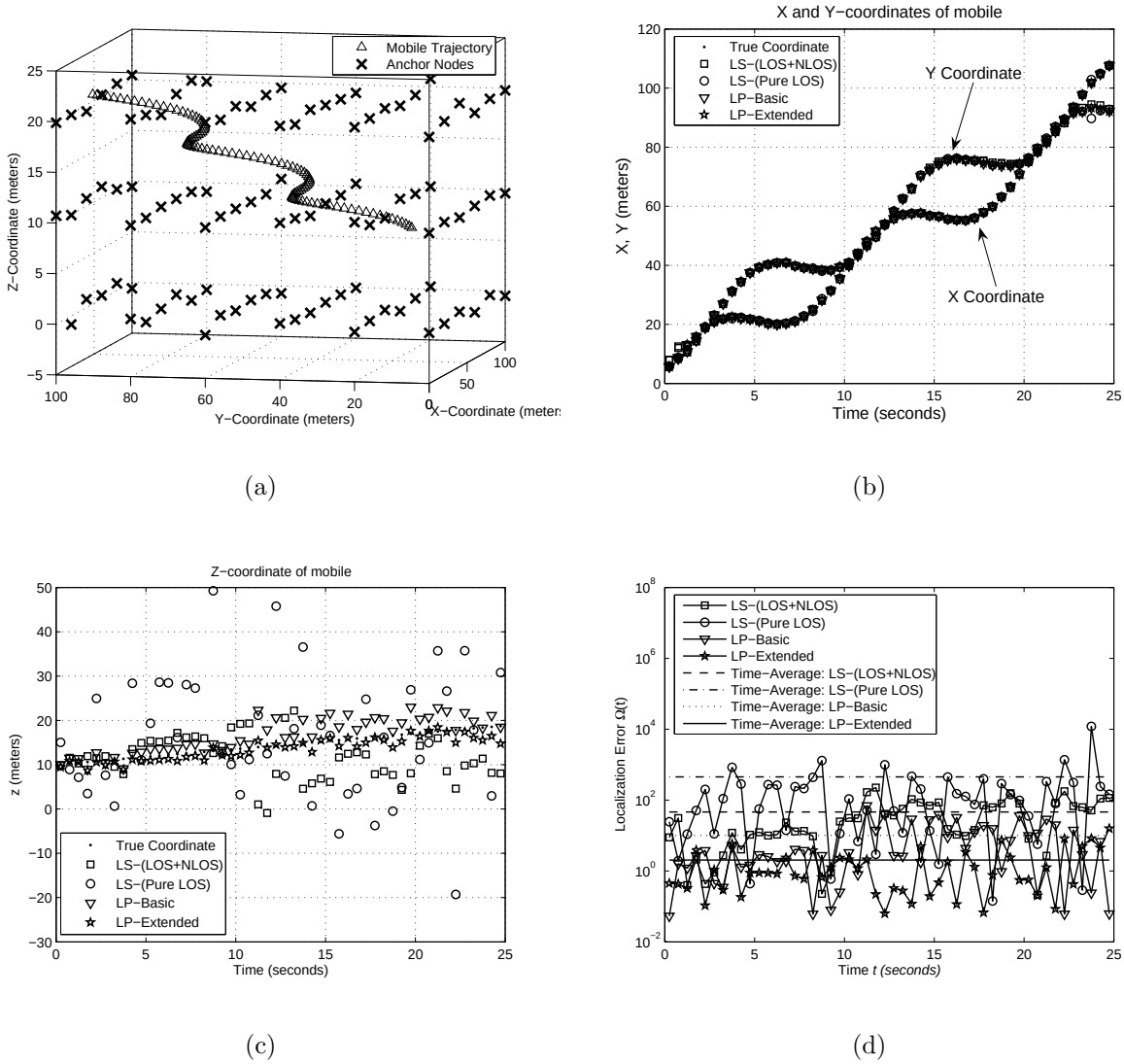


Figure 4.18: In this 3D indoor tracking simulation, $W = 50$, $H = 10$, $N_R = 108$, $\boldsymbol{x}_{M0} = [0\ 0\ 10]^T$ and $v = 1.5$ m/s. The remaining simulation parameters are the same as those in Figures 4.17(a) and 4.17(b). The time-axis is assumed to run from $t = 0$ to $t = 25$ seconds, with $T_s = 0.5$ second. **(a)** An illustration of the trajectory of the mobile node and the distribution of anchor nodes in a building. **(b)** The x and y components of the localization error $\Omega(t)$ versus time t. **(c)** The z component of the localization error $\Omega(t)$ versus time t. **(d)** The total localization error $\Omega(t)$ as a function of time t. The time-averaged localization error is also shown for comparison.

4.12 Appendix 4A: Distribution of $\hat{\tau}_0$ under NLOS conditions

Under NLOS propagation conditions, the TOA estimate is modeled as

$$\hat{\tau}_0 = \frac{d}{c} + n_\tau + b_\tau, \; H = H_1,$$

where n_τ is a zero-mean Gaussian random variable with variance σ_τ^2 and b_τ is exponentially distributed with mean λ_b. Therefore, the probability density function of $\hat{\tau}_0$ is defined by a convolution of the density functions of n_τ and b_τ:

$$\begin{aligned}
f_{\hat{\tau}_0}(\tau;\; d\,|H_1) &= \left(2\pi\lambda_b^2\sigma_\tau^2\right)^{-1/2} \int_0^\infty \exp\left(-\frac{(\tau - \frac{d}{c} - b)^2}{2\sigma_\tau^2}\right) \exp\left(-\frac{b}{\lambda_b}\right) db, \\
&= \left(\pi\lambda_b^2\right)^{-1/2} \int_{-\infty}^{\frac{\tau-\frac{d}{c}}{\sqrt{2}\sigma_\tau}} \exp\left(-y^2\right) \exp\left(-\frac{\tau - \frac{d}{c} - \sqrt{2}\sigma_\tau y}{\lambda_b}\right) dy, \\
&= \left(\pi\lambda_b^2\right)^{-1/2} \exp\left(\frac{\sigma_\tau^2}{2\lambda_b^2}\right) \exp\left(-\frac{\tau - \frac{d}{c}}{\lambda_b}\right) \int_{-\infty}^{\frac{\tau-\frac{d}{c}}{\sqrt{2}\sigma_\tau}} \exp\left(-\left(y - \frac{\sigma_\tau}{\sqrt{2}\lambda_b}\right)^2\right) dy, \\
&= \frac{1}{2\lambda_b} \exp\left(\frac{\sigma_\tau^2}{2\lambda_b^2}\right) \exp\left(-\frac{(\tau - \frac{d}{c})}{\lambda_b}\right) \left[1 + \operatorname{erf}\left(\frac{\lambda_b(\tau - \frac{d}{c}) - \sigma_\tau^2}{\sqrt{2}\sigma_\tau\lambda_b}\right)\right],
\end{aligned}$$

where $\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy$.

4.13 Appendix 4B: Derivation of $P_e^\star$ and $T^\star$ for positive, Gaussian RDS estimates

The probability density functions of the RDS estimates in LOS and NLOS scenarios are modeled as positive (truncated) Gaussian density functions:

$$f(x) = \begin{cases} \frac{1}{Q\left(-\frac{\mu_L}{\sigma_L}\right)} \cdot \frac{1}{\sqrt{2\pi\sigma_L^2}} e^{-\frac{(x-\mu_L)^2}{2\sigma_L^2}} u(x), & H = H_0, \\ \frac{1}{Q\left(-\frac{\mu_N}{\sigma_N}\right)} \cdot \frac{1}{\sqrt{2\pi\sigma_N^2}} e^{-\frac{(x-\mu_N)^2}{2\sigma_N^2}} u(x), & H = H_1, \end{cases} \tag{4.38}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-x^2/2}$, and $u(x)$ is the Heaviside function. The means μ_L, μ_N and the standard deviations σ_L, σ_N are specified in Table 4.2. Let the threshold selected for differentiating between LOS and NLOS RMS delay spread estimates be T. For a given

sample RMS delay spread measurement Z, the probability of error in the detection of the nature of the channel is given by

$$P_e(T) = p_L \cdot \Pr\{Z \geq T|H_0\} + p_N \cdot \Pr\{Z < T|H_1\}. \tag{4.39}$$

It must be pointed out for a given ad hoc indoor scenario, the probability p_L that a given link between two nodes is LOS is not necessarily equal to the probability p_N that the link is NLOS. In the absence of *a priori* knowledge, we assume $p_L = p_N = \frac{1}{2}$. From (4.38) and (4.39), we obtain

$$P_e(T) = p_L \cdot \frac{Q\left(\frac{T-\mu_L}{\sigma_L}\right)}{Q\left(-\frac{\mu_L}{\sigma_L}\right)} + p_N \cdot \left(1 - \frac{Q\left(\frac{T-\mu_N}{\sigma_N}\right)}{Q\left(-\frac{\mu_N}{\sigma_N}\right)}\right).$$

To find the optimal threshold $T^\star$ that minimizes the probability of error, we set $\frac{d}{dT}P_e\big|_{T=T^\star} = 0$, and it is straightforward to show that $T^\star$ is the solution of the quadratic equation

$$T^{\star 2}\left(\frac{1}{\sigma_L^2} - \frac{1}{\sigma_N^2}\right) + 2T^\star\left(\frac{\mu_N}{\sigma_N} - \frac{\mu_L}{\sigma_L}\right) + \frac{\mu_L^2}{\sigma_L^2} - \frac{\mu_N^2}{\sigma_N^2} = 2\log\left(\frac{p_L\sigma_N Q\left(-\frac{\mu_N}{\sigma_N}\right)}{p_N\sigma_L Q\left(-\frac{\mu_L}{\sigma_L}\right)}\right). \tag{4.40}$$

Then, minimum probability of channel detection error is given by $P_e^\star = P_e(T^\star)$.

4.14 Appendix 4C: Expression for q_j, $j = 1, 2, \cdots, m_N$

The term q_j is the probability that the true solution lies inside the jth circular constraint:

$$q_j = \Pr\{r_{Nj} > R_{Nj}\} = \Pr\{b_{Nj} + n_{Nj} > 0\},\ j = 1, 2, \cdots, m_N.$$

Assuming that $n_{Nj} \sim \mathcal{N}(0, \sigma_{Nj}^2)$ and $b_{Nj} \sim \mathcal{U}(0, B_{\max})$,

$$\begin{aligned}
q_j &= \int_0^{B_{\max}} \Pr\{n_{Nj} > -b\}\left(\frac{1}{B_{\max}}\right) db \\
&= \frac{1}{B_{\max}}\int_0^{B_{\max}}\left(\int_{-b}^{\infty}\frac{1}{\sqrt{2\pi\sigma_{Nj}^2}}\exp\left(-\frac{x^2}{2\sigma_{Nj}^2}\right)dx\right)db \\
&= 1 - \frac{1}{2B_{\max}}\int_0^{B_{\max}}\mathrm{erfc}\left(\frac{b}{\sqrt{2}\sigma_{Nj}}\right)db \\
&= 1 - \frac{1}{2}\mathrm{erfc}\left(\frac{z_j}{\sqrt{2}}\right) - \frac{1}{\sqrt{2\pi}}\left(\frac{1-e^{-\frac{z_j^2}{2}}}{z_j}\right),\ z_j = \frac{B_{\max}}{\sigma_{Nj}}.
\end{aligned}$$

Chapter 5

The Propagation of Localization Error and its Mitigation

5.1 Introduction

In chapter 1, we briefly discussed the classification of PoLoNets based on several criteria, such as network architecture, connectivity, and modes of location estimation. In infrastructure-based networks, anchor nodes are extensively deployed within the area of interest, allowing direct tracking of the locations of mobile nodes. In the ad hoc network architecture, anchor nodes could be deployed inside or outside the area of interest with connectivity to a subset of reference nodes. We then rely on reference nodes to propagate location information into the area of interest, thereby allowing mobile tracking. Further, unlocalized nodes (reference or mobile) may have full or partial connectivity with other unlocalized and anchor nodes in the network, which determines the number of range estimates that can be gathered.

In this chapter, we focus our attention on the structure of location-estimation in ad hoc PoLoNets. In infrastructure-based networks with dense deployment of anchors within the area of interest, reference nodes are unnecessary as mobile nodes can directly estimate their locations in a distributed manner, using range estimates from anchor nodes. However, in the ad hoc case with sparsely deployed anchors, either inside or outside the area of interest, three modes of location-estimation are possible: (i) fully distributed (when the requisite connectivity[1] with anchors is provisioned), (ii) centralized, and (iii) sequential. When all unlocalized nodes have connectivity with 3 or more anchors (for two-dimensional localization), location estimates can be generated in a fully distributed manner, i.e., location-estimates can be computed based only on *local* range and location information and without collaboration with other unlocalized nodes. The second approach is to route all the available range and

[1]In two dimensions, this implies that each unlocalized node has connectivity to at least three localized nodes when its location needs to be estimated.

Architecture	Connectivity	Estimation
Ad Hoc	Full	Distributed/Centralized/Sequential
Ad Hoc	Partial	Centralized/Sequential
Infrastructure-based	Full	Distributed
Infrastructure-based	Partial	Distributed/Centralized/Sequential

Table 5.1: Network Classifications and Options

location information, including range estimates between unlocalized nodes, to a centralized "solver" which jointly estimates the locations of all unlocalized nodes in the network and then relays the location estimates back to each node. The final approach is sequential estimation: nodes that obtain a sufficient number of range estimates from anchors can estimate their own locations in a distributed fashion, and then can act as anchors for other unlocalized nodes. Thus, location information is sequentially propagated from "true" anchors to the unlocalized nodes. It must be pointed out that hybrid modes of location estimation are possible for the different phases of localization. For example, the centralized or sequential approaches can be used for estimating the locations of reference nodes, whereas the locations of mobile nodes can be computed in a distributed manner thereafter.

Table 5.1 lists potential combinations of network architecture, extent of connectivity and the three modes of location-estimation discussed above. Each of the three modes of estimation mentioned above has certain advantages and disadvantages. In fully-distributed location-estimation, each unlocalized node relies only on range estimates from true anchors to obtain location estimates. A small number of anchors imposes strict connectivity requirements as any unlocalized node that does not have direct connectivity with 3 or more anchors cannot estimate its location. These requirements may not be feasible over large distances, especially in NLOS propagation environments, and thus, fully distributed methods are better suited to infrastructure-based networks with dense deployment of anchors within the area of interest. As we will see in this chapter, even if these connectivity constraints are satisfied, range estimation over large distances and the geometric dilution of precision associated with the anchor locations can result in poor localization accuracy. However, provided these connectivity requirements are satisfied, the complexity of implementing location-estimation and the communication cost are minimal, as no collaboration with other unlocalized nodes is required.

In the case of the centralized solver, instead of each node attempting to compute its location based on local range estimates, all range estimates and anchor locations are relayed to a node with superior computational power, which integrates *all* available range information to estimate the locations of all unlocalized nodes simultaneously. These location estimates are then routed back to the respective unlocalized nodes. Such a method has several advantages:

- All nodes are not required to be capable of estimating their locations. This allows for

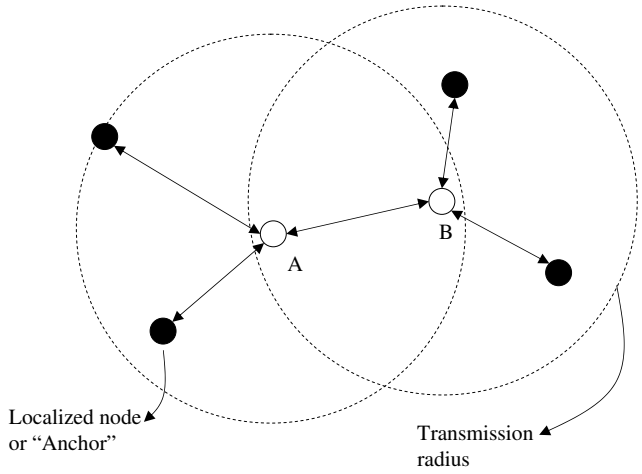


Figure 5.1: A scenario where the distributed computation of location-estimates is not feasible. Unlocalized nodes A and B cannot estimate their locations independently: although both nodes can obtain 3 range estimates, neither has 3 range estimates from *known* anchor locations.

the reduction in computational complexity within a majority of the nodes, as only the centralized solver performs large-scale computation.

- The availability of complete range information could potentially result in improved localization accuracy through the use of complex centralized algorithms. Algorithms can be devised which take advantage of range estimates between unlocalized nodes, which are not utilized in distributed estimation.

- In low-density networks, situations may arise where a distributed solution may not be feasible. Consider, for instance, the situation depicted in Figure 5.1: the unlocalized nodes A and B cannot estimate their locations independent of one another, even though each node can obtain 3 range estimates. In such scenarios, collaborative solutions are required, where locations of unlocalized nodes are jointly estimated using the available range and location information.

However, the centralized approach has several disadvantages as well:

- In the presence of mobility, which requires frequent updates of location-estimates, *all* range and location estimates would have to be routed to the centralized solver at regular intervals, which results in routing bottlenecks and higher interference levels [72]. Thus, such methods may be more suitable for static networks.

- In order to forward the location information to a central node, a route to the centralized solver must be known. Since the locations of unlocalized nodes are unknown *a priori*,

a routing protocol other than location-based routing is required. This approach therefore incurs additional communication cost that is dependent on the efficiency of the existing network protocols [73]. Further, in highly mobile applications and/or in dense networks, this may lead to considerable packet congestion and unacceptable latencies.

- The nodes around the centralized solver can become heavily congested with packets (and tend to consume a large amount energy when compared to other nodes) as range estimates to, and location estimates from, the centralized solver need to be routed through these nodes.
- Robustness is a key concern: if the routes to the centralized solver are broken (e.g., due to node failure), range and location information cannot be exchanged between the unlocalized nodes and the centralized solver.

It should be evident that the centralized approach is the most general mode of location-estimation, as the distributed localization techniques can also be implemented in a centralized manner. Due to the availability of additional range information between unlocalized nodes, an intuitive prediction is that the centralized approach should outperform the distributed approaches in terms of average localization accuracy. However, this gain in localization accuracy is realized at higher communications cost. The cost model devised in [74] predicts that the communication cost of the centralized approach is $\mathcal{O}(m\sqrt{N})$ higher than the fully distributed approach, where m is the number of range measurements per unlocalized node, and N is the number of nodes (assumed to be uniformly distributed over space) in the network.

The sequential location estimation approach appears to lie in between the fully-distributed and centralized approaches in terms of complexity and communication cost. Locations of nodes are estimated in a distributed manner; unlocalized nodes estimate their locations, and subsequently act as anchors for other unlocalized nodes. This can potentially achieve higher localization accuracy than the fully distributed approach, and lower communication cost than the centralized approach at the expense of higher computational complexity per node.

The main disadvantage of sequential location-estimation is (i) the inability to deal with degenerate situations such as those shown in Figure 5.1, and (ii) the propagation of localization error. The latter effect stems from the fact that several unlocalized nodes need to estimate their locations based on noisy range and location estimates from other previously unlocalized nodes, and thus, the errors in the location-estimates can accumulate. This implies that nodes farther away from true anchors are likely to have less accurate location estimates than nodes in the vicinity of true anchors. In the context of the fire-fighter tracking network, the propagation of localization error is a critical issue, as the errors in the propagated anchors can result in large errors in the fire-fighter's location estimate, even with accurate range estimates. Thus, there may be areas of the indoor environment (far away from the true

anchors) where mobile localization accuracy may be poor due to the presence of propagated anchors with inaccurate location estimates.

In this chapter, we investigate the impact of the propagation of localization error on the performance of sequential location-estimation schemes in ad hoc PoLoNets. We restrict our attention to the estimation of reference node locations, as accurate reference node location estimates imply accurate mobile location estimates, provided sufficient range measurement accuracy. We attempt to gain insight into the nature of the propagation of error in such networks, allowing the development of error mitigation schemes.

Previous Work included in Chapter 5: The CRLB for centralized estimation was derived in [19].

Key contributions included in this chapter are:

- The propagation of error problem and its mitigation in ad hoc sequential localization schemes has not been previously investigated except, to the best of our knowledge, in [75], [76]. A systematic study of the bounds on the performance of different estimation approaches, followed by the comparison of practical implementations of these approaches, is presented.
- Insights into the nature of the propagation of localization error, and its impact on the design of ad hoc PoLoNets are provided. These insights allow us to enumerate various means of limiting the propagation of localization error.
- The propagation of localization error in NLOS environments has not been studied previously. A novel method of mitigating the propagation of localization error, that utilizes range estimates over multiple hops while incorporating NLOS range estimates, is proposed.

This chapter is organized as follows: in Section 5.2, we formulate the CRLB on location-estimation for the fully-distributed and centralized estimation approaches. This quantifies the gains provided by the centralized approach in terms of average localization accuracy over the distributed approaches, and serves as a benchmark for the performance of sequential location estimation. Section 5.3 discusses the propagation of localization error in a simplistic one-dimensional localization scenario in order to gain fundamental insights into the nature of error propagation in the absence of variations in node geometry. An equivalent range model for sequential location-estimation is formulated to model anchor position errors, and an approximate CRLB for sequential location-estimation in two dimensions is derived based on this model in Section 5.3.2. The trends observed in the CRLB analysis are then compared with the trends observed with practical distributed and centralized methods in Section 5.4. We study the impact of geometry on the propagation of localization error, and the overall implications of the propagation of localization error on ad hoc PoLoNet design in Section

5.5. A novel multi-hop range-based method for mitigating the propagation of error that is applicable to NLOS environments is proposed in Section 5.6. Our conclusions are presented in Section 5.7.

5.2 CRLB for Fully Distributed and Centralized Location Estimation

In order to jointly formulate the distributed and centralized estimation approaches, we assume that there are N unlocalized nodes deployed within the area of interest, and M anchor nodes confined to a given region outside the area of interest. The unlocalized nodes are indexed $\{i = 1, 2, \cdots, N\}$, and the anchor nodes are indexed $\{i = N+1, N+2, \cdots, N+M\}$. The total number of nodes in the network is denoted by $N_t = N + M$. The ith unlocalized node is assumed to maintain connectivity with, and thereby receive range estimates from, a subset of the nodes in the network denoted by $\mathcal{T}_i$, $i = 1, 2, \cdots, N$. Since this set of nodes may comprise anchors *and* other unlocalized nodes, it is further partitioned into two subsets: $\mathcal{T}_i = \mathcal{A}_i \cup \mathcal{U}_i$, where $\mathcal{A}_i$ and $\mathcal{U}_i$ respectively denote the sets of anchor nodes and unlocalized nodes providing range estimates to node i.

We temporarily assume that all links between nodes in the network are LOS. This assumption is relaxed in Section 5.6. The range estimates available at the ith unlocalized node can then be modeled as:

$$r_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\| + n_{ij},\ i = 1, 2, \cdots, N, j \in \mathcal{T}_i.$$

where, n_{ij} is zero-mean Gaussian range measurement noise, with variance σ_{ij}^2. The range estimate variances are assumed to increase with the distance between the nodes, and are modeled as: $\sigma_{ij}^2 = K_E R_{ij}^2$, $R_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$. It is evident that for two-dimensional location-estimation, if $|\mathcal{A}_i| \geq 3$, each unlocalized node can estimate its location in a fully distributed manner, independent of other unlocalized nodes. However, if $|\mathcal{A}_i| < 3$, then the information from other unlocalized nodes needs to be incorporated.

5.2.1 Fully Distributed Solver

When $|\mathcal{A}_i| \geq 3$, $\forall i$, each unlocalized node can estimate its location based solely on range estimates from anchor nodes. When the information contained in $\mathcal{U}_i$ is ignored, the CRLB for the location-estimate of the ith unlocalized node is well-known [19], [9], and has also been derived in Chapter 4. In terms of the localization error $\Omega_{\mathbf{x},i}$ of the ith node,

$$\Omega_{\mathbf{x},i} \geq \left(\sum_{j \in \mathcal{A}_i} \frac{1}{\sigma_{ij}^2} \right) \left(\sum_{j \in \mathcal{A}_i} \sum_{k \in \mathcal{A}_i,\ k > j} \frac{\sin^2(\alpha_{ij} - \alpha_{ik})}{\sigma_{ij}^2 \sigma_{ik}^2} \right)^{-1}, \tag{5.1}$$

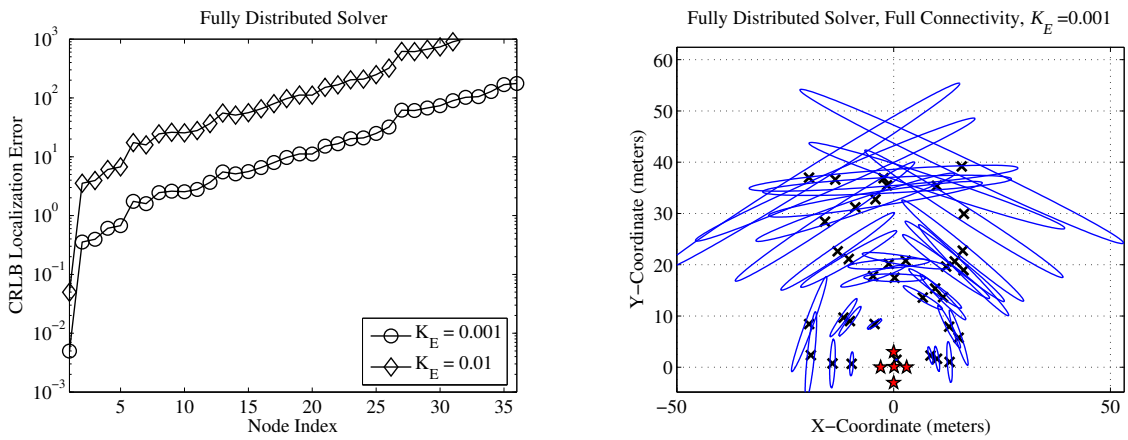
where, as before (see Chapter 4), $\alpha_{ij} = \measuredangle(\mathbf{x}_j - \mathbf{x}_i)$ is the orientation of anchor node j, $j \in \mathcal{A}_i$, relative to node i. The CRLB in terms of the covariance matrix of the location estimate $\hat{\mathbf{x}}_i$ is given by

$$\mathbf{C}_{\hat{\mathbf{x}}_\mathbf{i}} \geq \left(\begin{bmatrix} \sum_{j\in\mathcal{A}_i} \frac{\cos^2(\alpha_{ij})}{\sigma_{ij}^2} & \sum_{j\in\mathcal{A}_i} \frac{\cos(\alpha_{ij})\sin(\alpha_{ij})}{\sigma_{ij}^2} \\ \sum_{j\in\mathcal{A}_i} \frac{\cos(\alpha_{ij})\sin(\alpha_{ij})}{\sigma_{ij}^2} & \sum_{j\in\mathcal{A}_i} \frac{\sin^2(\alpha_{ij})}{\sigma_{ij}^2} \end{bmatrix} \right)^{-1}. \tag{5.2}$$

Figures 5.2(a) and 5.2(b) illustrate the CRLB on fully distributed location- estimation through the average localization error and error-ellipses respectively, with $M = 5$ anchor nodes located outside the area of interest, which is an $L \times L$ region, with $L = 50$ meters. The $N = 36$ unlocalized nodes are assumed to be uniformly (randomly) distributed over the area of interest and are sorted based on their increasing distance from the centroid of the anchors (the origin). Constant transmit power and full connectivity are assumed, i.e., each unlocalized node receives range estimates from all anchors.

As seen in Figures 5.2(a) and 5.2(b), the localization accuracy of the unlocalized nodes degrades with increasing distance from the centroid of anchor locations. This results from the increase in (i) the range estimate variances, and (ii) the geometric dilution of precision, as the distances from anchor nodes increase. From Figure 5.2(b), we observe that the scaling the range variances by a multiplicative factor (increasing K_E) also results in the corresponding scaling of the average localization error. We further note that the orientations of the error ellipses are determined by the geometrical orientations of anchor nodes relative to unlocalized nodes. Also noteworthy is the presence of "edge effects": the geometry of the anchor nodes relative to unlocalized nodes on the boundary of the region of interest typically result in larger localization errors than nodes in the interior of the area of interest (for a fixed distance from the centroid of the anchors).

The use of the fully distributed method described above may be inefficient due to constraints on the transmit power, as full connectivity between all unlocalized nodes of the network and all anchor nodes is required. Since the anchor nodes are assumed to be deployed over a small area, this implies that distant unlocalized nodes require large transmit powers to maintain connectivity. Further, for a fixed transmit power, as the distance between the unlocalized node and the anchors nodes increases, the range estimate variances and the geometrical dilution of precision increase, resulting in considerable degradation in localization accuracy. Therefore, from the standpoint of localization accuracy, the fully distributed approach is not well-suited to the ad hoc network architecture. Hence, we would prefer a centralized or sequential method where full connectivity is not required. The bounds on the performance of centralized location-estimation are the subject of the following subsection.



(a) $N = 36$, $M = 5$

(b) $N = 36$, $M = 5$

Figure 5.2: The area of interest is an $L \times L$ region, $L = 50$ meters. $N = 36$ unlocalized nodes are uniformly (randomly) distributed over this region. The unlocalized nodes, $i = 1, 2, \cdots, N$ are sorted with increasing distance from the centroid of anchors. (a) The localization error versus sorted node index, and (b) the 99% error ellipses corresponding to the fully distributed CRLB.

5.2.2 Centralized Solver

In order to quantify the gains of using a centralized solver in terms of localization accuracy relative to the fully distributed method described in the previous section, we first look at the CRLB on the localization error of unlocalized nodes. In the case of full connectivity, we assume that every *available* pair of range estimates between localized and unlocalized nodes is forwarded to the centralized solver, i.e., the set $\mathcal{R}$ of range estimates

$$\mathcal{R} = \{r_{ij} \,|r_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\| + n_{ij}\}\,,\ j \in \mathcal{T}_i,\ i = 1, 2, \cdots, N. \tag{5.3}$$

The goal is to estimates the vector of locations:

$$\mathbf{X} = \begin{bmatrix} x_1 & y_1 & x_2 & y_2 & \cdots & x_N & y_N \end{bmatrix}^T_{2N\times 1}, \tag{5.4}$$

given $\mathcal{R}$, and $\mathbf{x}_i$, $i = N+1, N+2, \cdots, N+M$. The CRLB for $\mathbf{X}$ has been derived in [19], and is also detailed in Appendix 5A for completeness. The CRLB can be expressed as:

$$\mathbf{C_X} \geq \mathbf{I}^{-1},$$

where the entries of the $(2N \times 2N)$ Fisher information matrix $\mathbf{I}$ are given by

$$
\begin{aligned}
I_{2i-1,2j-1} &= \begin{cases} \sum_{j \in \mathcal{T}_i} \frac{\cos^2 \alpha_{ij}}{\sigma_{ij}^2} & i = j, \\ -\frac{\cos^2 \alpha_{ij}}{\sigma_{ij}^2} & i \neq j,\ j \in \mathcal{U}_i, \\ 0 & \text{otherwise.} \end{cases} \\
I_{2i,2j} &= \begin{cases} \sum_{j \in \mathcal{T}_i} \frac{\sin^2 \alpha_{ij}}{\sigma_{ij}^2} & i = j, \\ -\frac{\sin^2 \alpha_{ij}}{\sigma_{ij}^2} & i \neq j,\ j \in \mathcal{U}_i, \\ 0 & \text{otherwise.} \end{cases} \\
I_{2i-1,2j} &= \begin{cases} \sum_{j \in \mathcal{T}_i} \frac{\cos \alpha_{ij} \sin \alpha_{ij}}{\sigma_{ij}^2} & i = j, \\ -\frac{\sin \alpha_{ij} \cos \alpha_{ij}}{\sigma_{ij}^2} & i \neq j,\ j \in \mathcal{U}_i, \\ 0 & \text{otherwise.} \end{cases} \\
&= I_{2i,2j-1} = I_{2j-1,2i} = I_{2j,2i-1}. \qquad (5.5)
\end{aligned}
$$

The total localization error for all unlocalized nodes is given by

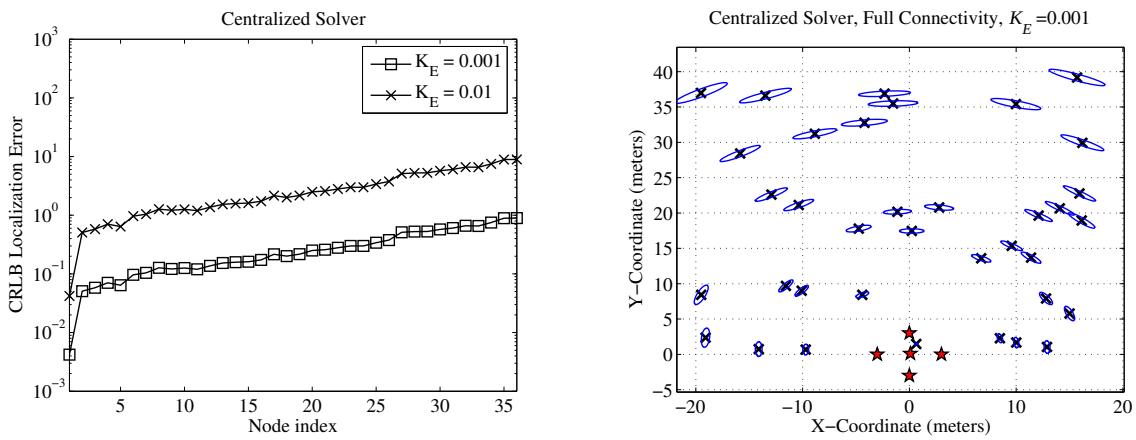
$$\Omega_t = \mathrm{Tr}\left(\mathbf{I}^{-1}\right).$$

The CRLB for the localization error of individual unlocalized nodes can also be extracted from the Fisher information matrix, as shown in Appendix 5A. The CRLB analysis [19] predicts that as the amount of range information increases, the localization error decreases. The CRLB for the centralized solver with full network connectivity is shown in Figure 5.3. In comparison to the fully distributed solver in Figure 5.2, we see that the availability of additional range information allows considerable reduction in the localization error. However, even in this case, we observe that as the distance to the anchors increases, the localization error also increases. This suggests that the increase in the localization error with distance from the anchors is inherent to the geometry of the ad hoc network architecture, and is the subject of Section 5.5.

In the following section, we discuss bounds on the performance of the sequential estimation approach, where the locations of unlocalized nodes are estimated sequentially, based on their distance from anchor nodes.

5.3 Distributed Sequential Location-Estimation

In the sequential mode of estimation, nodes that estimate their locations, in turn, serve as anchors for other unlocalized nodes. This approach does not require full connectivity, or connectivity with "true" anchors, and location-estimates can be obtained in a distributed manner. The main issue with a sequential localization approach is that using the location estimates of unlocalized nodes as anchor locations along with noisy range measurements can



(a) $N = 36$, $M = 5$

(b) $N = 36$, $M = 5$

Figure 5.3: The area of interest is an $L \times L$ region, $L = 50$ meters. $N = 36$ unlocalized nodes are uniformly (randomly) distributed over this region. The unlocalized nodes, $i = 1, 2, \cdots, N$ are sorted with increasing distance from the centroid of anchors. (a) The localization error versus sorted node index, and (b) the 99% error ellipses corresponding to the centralized CRLB.

result in the degradation of node localization accuracy. Further, these errors can propagate, as the distance between the unlocalized nodes and "true" anchors increases.

In this section, our goal is to quantify the extent of the propagation of error in two-dimensional sequential location-estimation, and gain insight into the steps which can be taken to mitigate this phenomenon. We begin by quantifying the propagation of error for the simplistic one-dimensional sequential location-estimation problem. The one-dimensional case is considered to side-step the influence of geometry of node locations, which makes the analysis considerably complicated, as will be seen in subsequent sections. We then extend this analysis in Section 5.3.2 to the two-dimensional location-estimation problem using an equivalent range estimate model to account for anchor position errors.

5.3.1 One-Dimensional Propagation of Error

In the one-dimensional location-estimation problem, we assume that unlocalized nodes are distributed along the positive x-axis, as shown in Figure 5.4. It is evident that each unlocalized node can estimate its location based on a single unbiased Gaussian range estimate for its predecessor node, and the corresponding node location estimate. For our purposes, we assume that the source of accurate position information is an anchor node at the origin or node 0. We first analyze the localization error versus distance, when an unlocalized node

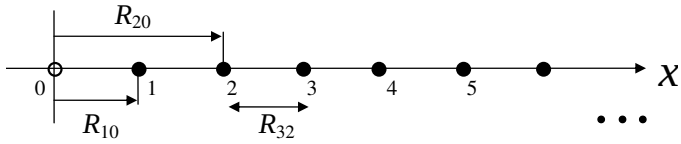


Figure 5.4: Propagation of localization error in along one dimension.

estimates its location solely based on a range estimate obtained from its preceding node.

Single range estimate from predecessor node

The kth unlocalized node receives a range estimate $r_{k,k-1}$ from its predecessor, node $(k-1)$, that is modeled as

$$r_{k,k-1} = R_{k,k-1} + w_{k,k-1},$$

where $R_{k,j} = x_k - x_j$, and $w_{k,k-1}$ is zero-mean Gaussian range measurement noise with variance $\sigma^2_{k,k-1} = K_E R^\beta_{k,k-1}$, β being the path loss exponent. Let $\hat{x}_i$ denote the estimate of the kth node's location. We have the following relations:

$$\begin{aligned}
\hat{x}_1 &= x_0 + r_{1,0} = r_{1,0} = R_{1,0} + w_{1,0} = x_1 + w_{1,0}, \\
\hat{x}_2 &= \hat{x}_1 + r_{2,1} = r_{1,0} + r_{2,1} = R_{1,0} + R_{2,1} + w_{1,0} + w_{2,1} = x_2 + w_{1,0} + w_{2,1}, \\
&\vdots \\
\hat{x}_k &= x_k + \sum_{j=1}^{k} w_{j,j-1}.
\end{aligned}$$

We see that the error in the kth node's location-estimate is a superposition of the range measurement errors in *all* previous range measurements. Assuming that the errors in the range estimates are independent, the variance of the kth location-estimate is given by:

$$\Omega_k = \sum_{j=1}^{k} \sigma^2_{j,j-1} = K_E \sum_{j=1}^{k} R^\beta_{j,j-1}.$$

If the nodes are equally spaced: $R_{j,j-1} = R$, $\forall j$, then

$$\Omega_k = k K_E R^\beta,$$

which implies that the localization error grows linearly with distance from the origin. Further, the extent of the increase in the localization error can be reduced by providing more accurate range estimates (smaller K_E) and a higher density of nodes (resulting in smaller inter-node distances R). If these options are not feasible, we can allow connectivity between each unlocalized node and several preceding nodes, thereby providing a larger number of range estimates for location-estimation. This can be achieved, for instance, by using a higher transmit power, and is discussed in the following subsection.

Multiple range estimates, equidistant nodes

Given range estimates from m preceding nodes, the kth node can estimate its location using the relation:

$$\hat{x}_k = \frac{1}{m}\sum_{j=1}^{m}\left(\hat{x}_{k-j} + r_{k,j}\right).$$

From the above relation, it can be shown that

$$\hat{x}_k = x_k + v_k,$$

where

$$v_k = \frac{1}{m}\sum_{j=1}^{m} v_{k-j} + y_k,\ \ y_k = \frac{1}{m}\sum_{j=1}^{m} w_{k,k-j}. \tag{5.6}$$

In order to provide connectivity with m preceding nodes, the transmit power has to be increased by a factor m^β, which (in the absence of network self-interference) also reduces the range estimate variances. We assume, as before, that $w_{k,k-j}$ are independent and $R_{j,j-1} = R$, $\forall j$. Then, $w_{k,k-j} \sim \mathcal{N}\left(0, K_E\left(\frac{jR}{m}\right)^\beta\right)$, and therefore

$$\sigma^2_{y_k} = \frac{K'}{m^2}\sum_{j=1}^{m}\left(\frac{j}{m}\right)^\beta, \tag{5.7}$$

where $K' = K_E R^\beta$. When $\beta = 2$, we have

$$\sigma^2_{y_k} = \sigma^2_y = \frac{K'(m+1)(2m+1)}{6m^3}, \quad \forall k. \tag{5.8}$$

We observe that the variances of y_k are independent of k. From (5.6), taking the z-transform of v_k, we have

$$\begin{aligned} V(z) &= \frac{V(z)}{m}\sum_{j=1}^{m} z^{-j} + Y(z) \Rightarrow V(z) = \frac{Y(z)}{\left(1 - \frac{1}{m}\sum_{j=1}^{m} z^{-j}\right)}, \qquad (5.9)\\ &= \begin{cases} \frac{Y(z)}{(1-z^{-1})\prod_{j=1}^{n_p}(1-z_{pj}z^{-1})\left(1-z^\star_{pj}z^{-1}\right)}, & m \text{ odd},\\ \frac{Y(z)}{(1-z^{-1}(1-r_0z^{-1})\prod_{j=1}^{n_p}(1-z_{pj}z^{-1})\left(1-z^\star_{pj}z^{-1}\right)}, & m \text{ even}, \end{cases}\\ &= \begin{cases} Y(z)\left[\frac{B}{1-z^{-1}} + \sum_{j=1}^{n_p}\left(\frac{A_j}{1-z_{pj}z^{-1}} + \frac{A_j^*}{1-z^*_{pj}z^{-1}}\right)\right], & m \text{ odd},\\ Y(z)\left[\frac{B}{1-z^{-1}} + \frac{A_0}{1-r_0z^{-1}} + \sum_{j=1}^{n_p}\left(\frac{A_j}{1-z_{pj}z^{-1}} + \frac{A_j^*}{1-z^*_{pj}z^{-1}}\right)\right], & m \text{ even}, \end{cases} \end{aligned}$$

where $|r_0|$, $|z_{pj}| < 1$, $\forall j$. Therefore, if $z_{pj} = r_j e^{\sqrt{-1}\phi_j}$, then

$$v_k = \begin{cases} y_k \star \left[B + \sum_{j=1}^{n_p} 2\Re(A_j^k)\cos\left(k\phi_j\right)\right] u_k, & m \text{ odd},\\ y_k \star \left[B + A_0 r_0^k + \sum_{j=1}^{n_p} 2\Re(A_j) r_j^k \cos\left(k\phi_j\right)\right] u_k, & m \text{ even}. \end{cases} \tag{5.10}$$

We can now obtain individual expressions for different values of m.

<u>Case 1: $m = 2$</u> Applying (5.9) with $m = 2$, we obtain

$$\begin{aligned} V(z) &= \frac{Y(z)}{1 - \frac{z^{-1}}{2} - \frac{z^{-2}}{2}} = Y(z)\left[\frac{2}{3\left(1 - z^{-1}\right)} + \frac{1}{3\left(1 + \frac{1}{2}z^{-1}\right)}\right], \\ \Rightarrow\ v_k &= y_k \star \left[\frac{1}{3}\left(2 + \left(-\frac{1}{2}\right)^k\right) u_k\right] = \frac{2}{3}\sum_{l=0}^{k-1} y_{k-l}\left(1 - \left(-\frac{1}{2}\right)^l\right). \end{aligned}$$

Therefore, the localization error of the kth node is given by

$$\Omega_k = E\left[v_k^2\right] = \frac{4\sigma_y^2}{9}\sum_{l=0}^{k-1}\left(1 - \left(-\frac{1}{2}\right)^l\right)^2 = \frac{5K'}{36}\left[k + 1 - \frac{2^{-k+1}}{3} - \frac{4^{-k}}{3}\right].$$

We see that when $k \gg 1$, $\Omega_k \approx \frac{5K'}{36}k$, indicating a linear increase in the localization error with the number of nodes.

<u>Case 2: $m = 3$</u> Applying (5.9) with $m = 3$, we obtain

$$V(z) = \frac{Y(z)}{\left(1 - z^{-1}\right)\left(1 - z_0 z^{-1}\right)\left(1 - z_0^\star z^{-1}\right)},$$

where $z_0 = \frac{-1+\sqrt{-2}}{3} = a_0 e^{j\phi_0}$. Therefore,

$$\begin{aligned} V(z) &= Y(z)\left[\frac{1}{2(1 - z^{-1})} + \frac{1}{4(1 - z_0 z^{-1})} + \frac{1}{4(1 - z_0^\star z^{-1})}\right] \\ \Rightarrow v_k &= y_k \star \left[\frac{1}{2} + \frac{1}{4}\left(z_0^k + z_0^{\star k}\right)\right] u_k = \frac{1}{2}\sum_{l=0}^{k-1} y_{k-l}\left(1 + a_0^l \cos\left(l\phi_0\right)\right). \end{aligned}$$

Therefore, the localization error of the kth node is given by

$$\Omega_k = \frac{1}{4}\sum_{l=0}^{k-1}\sigma_{y_{k-l}}^2\left(1 + a_0^l\cos\left(l\phi_0\right)\right)^2 = \frac{7K'}{162}\left[k + \sum_{l=0}^{k-1}\left(2a_0^l\cos\left(l\phi_0\right) + a_0^{2l}\cos^2\left(l\phi_0\right)\right)\right].$$

It can be verified through the above expression, that Ω_k increases approximately linearly with k when $k \gg 1$, albeit at a lower rate than when $m = 2$. Proceeding in the above manner, the localization error can be derived for all values of m. In the following subsection, we analyze the localization error when range estimates from all predecessors (including the anchor node at the origin) are used to estimate the kth node's location.

Using range estimates from *all* predecessors, equidistant nodes:

When range estimates from all predecessors are used, we have $m = k$, and the following relations:

$$\begin{aligned}
\hat{x}_k &= x_k + v_k, \text{ where} \\
v_1 &= y_1, \\
v_2 &= \frac{1}{2}(x_0 + v_1) + y_2 = \frac{y_1}{2} + y_2, \\
v_3 &= \frac{1}{3}(v_0 + v_1 + v_2) + y_3 = \frac{1}{3}\left(y_1 + \frac{y_1}{2} + y_2\right) + y_3 = \frac{y_1}{2} + \frac{y_2}{3} + y_3, \\
v_4 &= \frac{1}{4}(v_0 + v_1 + v_2 + v_4) + y_4 = \frac{y_1}{2} + \frac{y_2}{3} + \frac{y_3}{4} + y_4.
\end{aligned}$$

Proceeding in this manner, we see that

$$v_k = y_k + \sum_{i=1}^{k-1} \frac{y_i}{i+1}, \; y_k = \frac{1}{k}\sum_{j=1}^{k} w_{k,k-j}. \tag{5.11}$$

Once again, $w_{k,k-j}$ are assumed to be independent with variances

$$\sigma^2_{k,k-i} = K'\left(\frac{i}{k}\right)^{\beta}, \; K' = K_E R^{\beta}.$$

Therefore,

$$\sigma^2_{y_k} = \frac{K'}{k^{2+\beta}} \sum_{j=1}^{k} i^{\beta},$$

and when $\beta = 2$,

$$\sigma^2_{y_k} = \frac{K'(k+1)(2k+1)}{6k^3}. \tag{5.12}$$

From (5.11), the localization error of the kth node is given by

$$\begin{aligned}
\Omega_k &= E\left[v_k^2\right] = \sigma^2_{y_k} + \sum_{i=1}^{k-1} \frac{\sigma^2_{y_i}}{(i+1)^2} = \frac{K'(k+1)(2k+1)}{6k^3} + \sum_{i=1}^{k-1} \frac{K'(2i+1)}{6i^3(i+1)} \\
&= \frac{K'}{6}\left[\frac{2}{k} + \frac{3}{k^2} + \frac{1}{k^3} + \sum_{i=1}^{k-1} \frac{2}{i^3} - \sum_{i=1}^{k-1} \frac{1}{i^3(i+1)}\right].
\end{aligned}$$

It can be verified that as $k \to \infty$, $\Omega_k \to 0.3K'$, i.e., the localization error does not diverge as $k \to \infty$. This suggests that providing a progressively larger number of range estimates to unlocalized nodes as their distance to the anchor (origin) increases offsets the propagation of error.

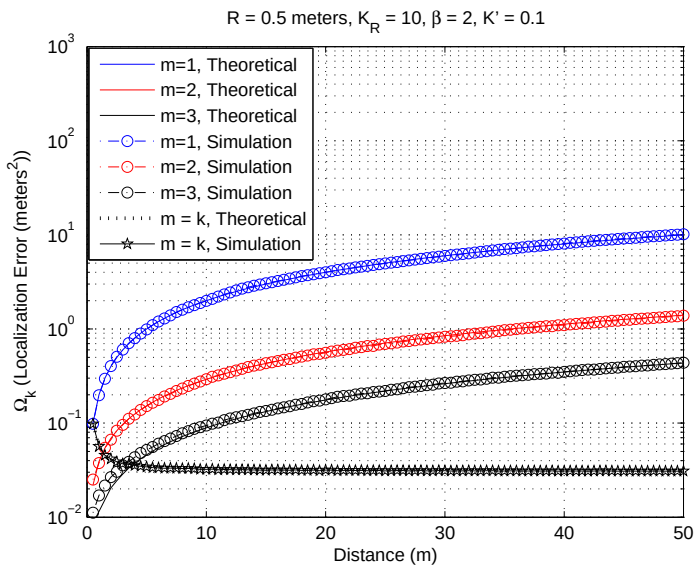


Figure 5.5: Variation of the localization error with distance when $K' = 0.1$. It can be shown that in cases with $m = \{1, 2, 3\}$, as $k \to \infty$, $\Omega_k \to \infty$. However, when $m = k$, as $k \to \infty$, $\Omega_k \to 0.3K'$.

To summarize the expressions for the localization error versus distance obtained in each of the discussed sub-cases:

$$\Omega_k = \begin{cases} \frac{5K'}{36}\left[k + 1 - \frac{2^{-k+1}}{3} - \frac{4^{-k}}{3}\right], & m = 2, \\ \frac{7K'}{162}\left[k + \sum_{l=0}^{k-1}\left(2\left(\frac{1}{3}\right)^{\frac{l}{2}}\cos\left(l\phi_0\right) + \left(\frac{1}{3}\right)^{l}\cos^2\left(l\phi_0\right)\right)\right], & m = 3, \\ \frac{K'}{6}\left[\frac{2}{k} + \frac{3}{k^2} + \frac{1}{k^3} + \sum_{i=1}^{k-1}\frac{2}{i^3} - \sum_{i=1}^{k-1}\frac{1}{i^3(i+1)}\right], & m = k. \end{cases} \tag{5.13}$$

The analytical expressions (based on (5.13)) and simulated variation of the localization error with distance from the origin for different values of m are shown in Figure 5.5. As seen in Figure 5.5, if all the available range estimates are utilized, there is no propagation of error and the localization error decreases with increasing distance from the reliable source of location information (the origin). However, this reduction in the localization error comes at the price of increasing transmit power as the distance from the origin increases, i.e., infinite transmit power as $k \to \infty$.

Therefore, if there are no constraints on the transmit power, the propagation of localization error can be completely offset by using a progressively larger number of range estimates for location estimation. However, for a fixed maximum transmit power, the localization error propagates, i.e., increases as the distance to the reliable source of location information (the origin) increases. This analysis suggests that in the presence of transmit power constraints that the distance over which the desired localization accuracy can be guaranteed may be

limited. In the analogous two-dimensional case, this would suggest that the area over which the desired average localization accuracy can be guaranteed may be bounded from above.

The above one-dimensional analysis further suggests that if the transmit power is fixed, then the propagation of localization error can be contained by (i) increasing node density (this increases m and reduces inter-node distances R, which in turn reduces range estimate variances), (ii) increasing range estimation accuracy (this can be achieved by increased range averaging, and reduces the value of K_E). Another key aspect of the above analysis that needs to be emphasized is that in the discussed one-dimensional location estimation scenario, the location estimates of nodes are unbiased and Gaussian distributed, due to the structure of the location estimator, and the fact that the range estimates are assumed to be unbiased Gaussian random variables. However, if the location estimates are biased, even if progressively larger number of range estimates are used to estimate the locations of nodes, the localization error can diverge as $k \to \infty$. For instance, if a positive bias b is included in each location estimate, the localization error diverges as $k \to \infty$. Thus, a key premise to the decrease in the sequential localization error with an increasing number of range estimates is that the available range and location estimates are unbiased. However, in the two-dimensional case, we saw in Section 3.6, that even if the available range estimates are unbiased and Gaussian distributed, the two-dimensional location estimates can be neither unbiased nor Gaussian distributed. Further, this suggests that the propagation of localization error will be exacerbated in NLOS propagation environments.

In the following section, we extend the analysis of the propagation of error to the two-dimensional location-estimation problem. The derivation of the CRLB for the accuracy of location estimates in the two-dimensional case, in the presence of both anchor position errors and range measurement errors, is known to be a hard problem [34], [25]. In order to analyze sequential localization in two-dimensions, we first derive an approximate CRLB for the case of localization with Gaussian anchor position errors.

5.3.2 Equivalent range measurement model with Gaussian anchor position errors

In order to characterize the impact of anchor location errors on two-dimensional localization accuracy, we derive the equivalent range measurement error that results from a given anchor position error. Consider an unlocalized node with unbiased Gaussian range estimates from localized nodes whose location-estimates are also noisy. Let $R = R(x_A, y_A) = \sqrt{(x - x_A)^2 + (y - y_A)^2}$ be the distance between an unlocalized node located at (x, y) and an anchor located at (x_A, y_A). In order to analyze the impact of anchor position errors, we temporarily neglect the presence of range measurement noise, i.e., the range estimate $r = R$. For small anchor position errors, Δx and Δy respectively in the x and y coordinates, we can

write the *effective range error* as

$$\Delta r \approx \left(\frac{\partial R}{\partial x_A}\right)\Delta x + \left(\frac{\partial R}{\partial y_A}\right)\Delta y = \Delta x \cos\alpha_A + \Delta y \sin\alpha_A, \tag{5.14}$$

where $\alpha = \arctan\left(\frac{y-y_A}{x-x_A}\right)$ represents the orientation of the anchor relative to x. Assuming that Δx and Δy are zero-mean Gaussian with joint covariance matrix $\mathbf{C}_\Delta$ given by

$$\mathbf{C}_\Delta = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix}, \tag{5.15}$$

we see from (5.14) that Δr is a zero-mean Gaussian random variable. The variance of Δr is given by

$$\begin{aligned} \sigma_r^2 &= E\left\{(\Delta r)^2\right\} = E\left\{(\Delta r)(\Delta r)^T\right\} \\ &= E\left\{[\cos\alpha_A \;\; \sin\alpha_A]\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}[\Delta x \;\; \Delta y]\begin{bmatrix} \cos\alpha_A \\ \sin\alpha_A \end{bmatrix}\right\} = \mathbf{e}_A^T \mathbf{C}_\Delta \mathbf{e}_A, \end{aligned} \tag{5.16}$$

where $\mathbf{e}_A = \begin{bmatrix} \cos\alpha_A \\ \sin\alpha_A \end{bmatrix}$. Substituting (5.15) in (5.16), we obtain

$$\begin{aligned} \sigma_r^2 &= [\cos\alpha_A \;\; \sin\alpha_A]\,\mathbf{C}_\Delta \begin{bmatrix} \cos\alpha_A \\ \sin\alpha_A \end{bmatrix} \\ &= [\cos\alpha_A \;\; \sin\alpha_A]\begin{bmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{yx} & \sigma_y^2 \end{bmatrix}\begin{bmatrix} \cos\alpha_A \\ \sin\alpha_A \end{bmatrix} \\ &= \left[\sigma_x^2\cos\alpha_A + \sigma_{yx}\sin\alpha_A \;\;\; \sigma_{xy}\cos\alpha_A + \sigma_y^2\sin\alpha_A\right]\begin{bmatrix} \cos\alpha_A \\ \sin\alpha_A \end{bmatrix} \\ &= \sigma_x^2\cos^2\alpha_A + 2\sigma_{xy}\sin\alpha_A\cos\alpha_A + \sigma_y^2\sin^2\alpha_A. \end{aligned}$$

It is a well known result [77] that for a square matrix $\mathbf{A}$ and for all vectors $\mathbf{x}$ of appropriate dimension,

$$\lambda_{\min}(\mathbf{A}) \le \frac{\mathbf{x}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}} \le \lambda_{\max}(\mathbf{A}), \tag{5.17}$$

where $\lambda_{\min}(\mathbf{A})$ and $\lambda_{\max}(\mathbf{A})$ are the minimum and maximum eigenvalues of $\mathbf{A}$. The values of $\mathbf{x}$ at which the equalities are achieved are the eigenvectors $\mathbf{x}_{\min}(\mathbf{A})$ and $\mathbf{x}_{\max}(\mathbf{A})$ corresponding to the minimum and maximum eigenvalues respectively. Since $\mathbf{e}_A^T\mathbf{e}_A = 1$ for all α_A, we have

$$\lambda_{\min}(\mathbf{C}_\mathbf{x}) \le \mathbf{e}_A^T\mathbf{C}_\mathbf{x}\mathbf{e}_A \le \lambda_{\max}(\mathbf{C}_\mathbf{x}), \tag{5.18}$$

and the equalities are satisfied when $\mathbf{e}_A = \frac{\mathbf{x}_{\min}(\mathbf{C}_\mathbf{x})}{\|\mathbf{x}_{\min}(\mathbf{C}_\mathbf{x})\|}$ and $\mathbf{e}_A = \frac{\mathbf{x}_{\max}(\mathbf{C}_\mathbf{x})}{\|\mathbf{x}_{\max}(\mathbf{C}_\mathbf{x})\|}$ respectively. In other words, the variance of the equivalent range error is maximized when the major axis of the error ellipse of the anchor location-estimate is aligned with the vector from the unlocalized node to the anchor node as seen in Figure 5.6(a). The variance of the equivalent range

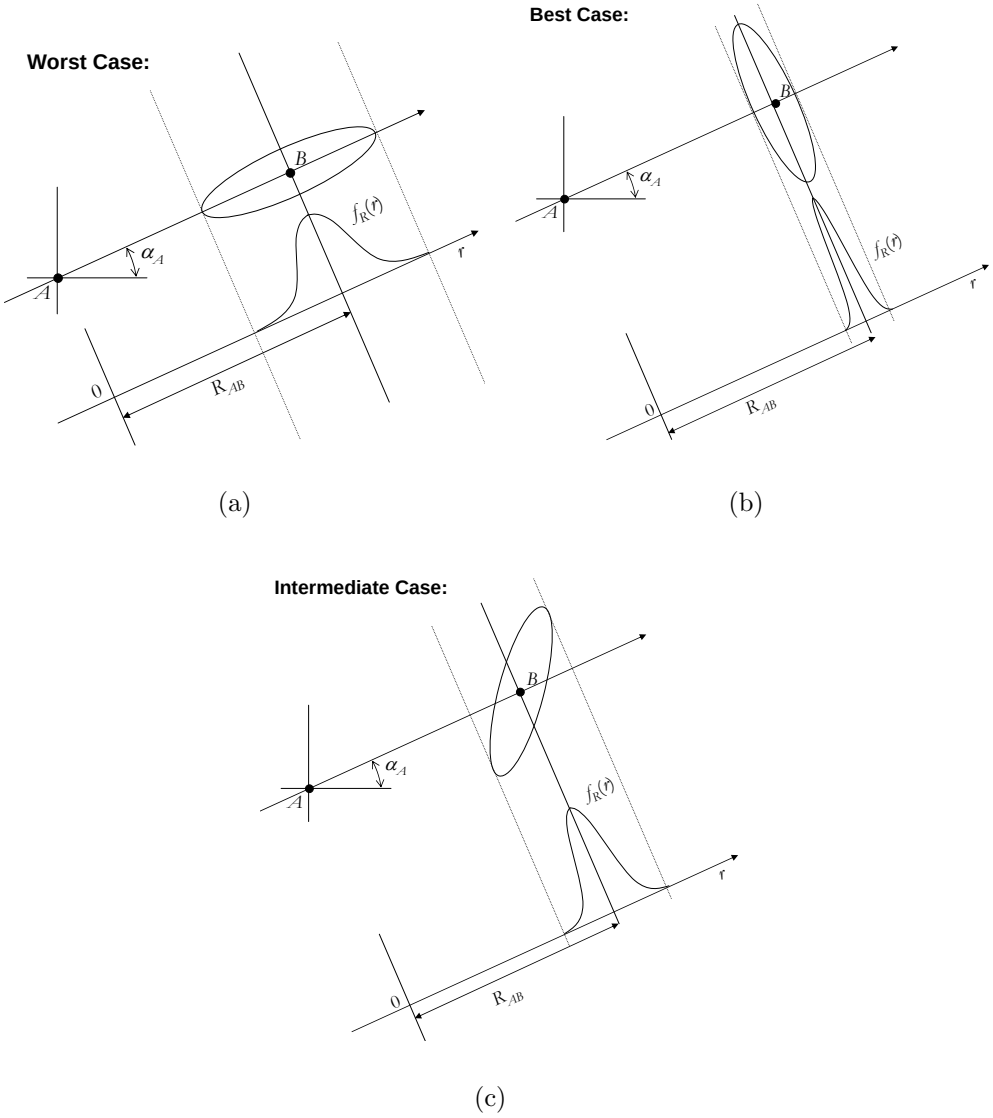


Figure 5.6: Impact of the two-dimensional distribution of Gaussian anchor position error on the distribution of the equivalent range error in the absence of range measurement noise, with $\sigma_{xy} \neq 0$. In (a), the vector $\mathbf{e}_A$ is the same as the eigenvector (normalized to unit norm) of the matrix $\mathbf{C_x}$ corresponding to the maximum eigenvalue. In (b), $\mathbf{e}_A$ is the same as the eigenvector (normalized to unit norm) of the matrix $\mathbf{C_x}$ corresponding to the minimum eigenvalue. In (c), $\mathbf{e}_A$ is a linear combination of the eigenvectors (normalized to unit norm) of the matrix $\mathbf{C_x}$ corresponding to the minimum and maximum eigenvalues.

error is minimized when the minor axis of the error ellipse is aligned with the vector from the unlocalized node to the anchor node, as see in Figure 5.6(b). An intermediate case is illustrated in Figure 5.6(c). If the position errors in the x and y coordinates are uncorrelated ($\sigma_{xy} = 0$), and identically distributed as $\mathcal{N}(0, \sigma^2)$, then the error ellipse reduces to a circle and it is easy to see that

$$\sigma_r^2 = \sigma^2, \tag{5.19}$$

irrespective of the orientation of the anchor node relative to the unlocalized node.

Thus, the effect of small (Δx, $\Delta y \ll R$) Gaussian anchor position errors can be modeled as an enhancement of the Gaussian range measurement noise. Further, in the presence of range measurement noise and anchor position errors, the approximate model for the range estimates becomes

$$r_i \sim \mathcal{N}(0, \sigma_i^2 + \sigma_{ri}^2), \tag{5.20}$$

where σ_i^2 denotes the variance of the range measurement noise, and σ_{ri}^2 denotes the equivalent range measurement noise attributed to anchor position errors, given by (5.16). For ease of notation, we use $\rho_i^2 \triangleq \sigma_i^2 + \sigma_{ri}^2$ as the total equivalent range estimate variance in the subsequent discussion.

Theorem 6 (Approximate CRLB with Gaussian Range and Anchor Position Errors). *The CRLB for the estimation of a location* $\mathbf{x}$ *of a node given* m *noisy Gaussian ranges* $r_i \sim \mathcal{N}(R_i, \sigma_i^2)$*, from anchors with* exactly *known locations* $\mathbf{x}_i$*,* $i = 1, 2, \cdots, m$*, is given by*

$$\mathbf{C}_{\mathbf{x}-\hat{\mathbf{x}}} \geq \mathbf{I}^{-1},$$

where the Fisher information matrix $\mathbf{I}$ *is given by*

$$\mathbf{I} = \begin{bmatrix} \sum_{i=1}^m \frac{\cos^2(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^m \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} \\ \sum_{i=1}^m \frac{\cos(\alpha_i)\sin(\alpha_i)}{\sigma_i^2} & \sum_{i=1}^m \frac{\sin^2(\alpha_i)}{\sigma_i^2} \end{bmatrix}. \tag{5.21}$$

This bound can be expressed in terms of the localization error

$$\begin{aligned} \Omega_x &= E\left\{(x-\hat{x})^2 + (y-\hat{y})^2\right\} = \sigma_x^2 + \sigma_y^2 \\ &\geq Tr\left(\mathbf{I}^{-1}\right) = \frac{\sum_{i=1}^m \frac{1}{\sigma_i^2}}{\sum_{i=1}^m \sum_{j=1, j>i}^m \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2 \sigma_j^2}}. \end{aligned} \tag{5.22}$$

Using the equivalent range model for anchor position errors, we can provide an "approximate" CRLB in the presence of anchor position errors. Assuming the errors in the position estimates of the anchors are zero-mean Gaussian distributed: $(\mathbf{x}_i - \hat{\mathbf{x}}_i) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_i)$*, then the approximate CRLB in terms of the localization error can be written as:*

$$\begin{aligned} \tilde{\Omega}_x &= E\left\{(x-\hat{x})^2 + (y-\hat{y})^2\right\} = \sigma_x^2 + \sigma_y^2 \\ &\gtrsim \frac{\sum_{i=1}^m \frac{1}{\rho_i^2}}{\sum_{i=1}^m \sum_{j=1, j>i}^m \frac{\sin^2(\alpha_i - \alpha_j)}{\rho_i^2 \rho_j^2}}. \end{aligned} \tag{5.23}$$

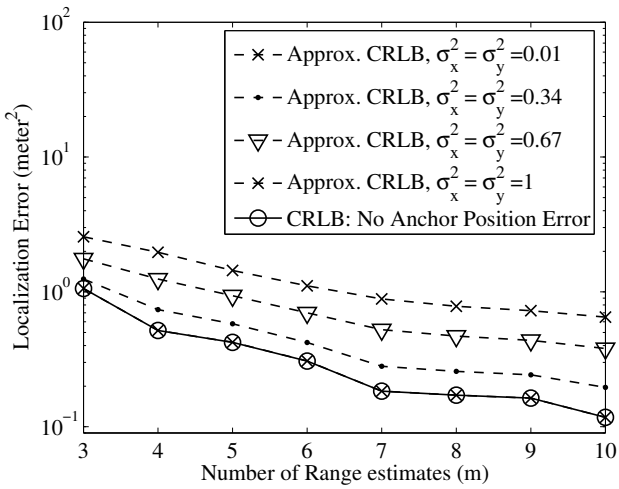


Figure 5.7: The approximate CRLB with anchor position errors, versus the number of range estimates. The position error of all the anchors are assumed to be i.i.d. zero-mean Gaussian, with equal variances in x and y coordinates $\sigma_x^2 = \sigma_y^2$ (meter^2, $\sigma_{xy} = 0$. The range estimate variances are obtained according to $\sigma_i^2 = K_E R_i^\beta$, with $K_E = 0.01$ and $\beta = 2$.

where, from (5.16) and (5.20),

$$\rho_i^2 = \sigma_i^2 + \mathbf{e}_i^T \mathbf{C}_i \mathbf{e}_i, \tag{5.24}$$

and $\mathbf{e}_i = [\cos\alpha_i \;\; \sin\alpha_i]^T$.

Figure 5.7 compares the CRLB with and without Gaussian anchor position errors in the presence of Gaussian range measurement errors. It must be emphasized that the above bound is applicable only for small anchor Gaussian position errors, i.e., $\Delta x, \Delta y \ll R$.

Given a set of N unlocalized nodes and M fixed anchors, assuming sufficient connectivity between anchors and unlocalized nodes to obtain location estimates in a distributed manner, the following sequential localization scheme for the unlocalized nodes can be applied. Let the true locations of the unlocalized nodes be denoted by $\mathbf{x}_i$, $i = 1, 2, \cdots, N$, and the locations of anchors by $\mathbf{x}_i$, $i = N+1, N+2, \cdots, N+M$. At stage k, let $\mathcal{A}_k$ denote the set of all nodes whose locations are either known (true anchors) or estimated. We assume that the unlocalized nodes $k = 1, 2, \cdots, N$, are arranged in the sequence of their location-estimation. Node $k = 1$ estimates its location based only on range estimates from the set of fixed anchors $\mathcal{A}$. Once node $k = 1$ estimates its location, it is added to the set of localized nodes. Therefore, the set of localized nodes $\mathcal{A}_2$ that provide range estimates to node $k = 2$ is governed by:

$$\mathcal{A}_2 \subset \mathcal{A} \cup \{\hat{\mathbf{x}}_1\}. \tag{5.25}$$

Node $k = 2$ then estimates its location based on range estimates from the nodes in $\mathcal{A}_2$. In general, node k, $k = 1, 2, \cdots, N$, estimates its location based on range estimates from a

subset of the set of localized nodes:

$$\mathcal{A}_k \subset \mathcal{A} \cup \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \cdots, \hat{\mathbf{x}}_{k-1}\}. \tag{5.26}$$

It is clear that $|\mathcal{A}_k| \leq |\mathcal{A}| + k - 1$. Since estimated locations of $(k-1)$ nodes contain estimation errors, this affects the estimation of the kth node's location. In order to model the estimation error at the kth stage, we resort to the following theorem.

Theorem 7 (Sequential Estimation with Range and Anchor Position Errors). *Let $\mathbf{C}_l$, $l = 1, 2, \cdots, k-1$, denote the covariance matrices of the position errors of the unlocalized nodes whose locations have been estimated. For true anchors, it is evident that the covariance matrices of the position errors is $\mathbf{0}_{2\times 2}$. Then we define*

$$\rho_{kl}^2 = \sigma_{kl}^2 + \mathbf{e}_{kl}^T \mathbf{C}_l \mathbf{e}_{kl}, \quad l \in |\mathcal{A}_k|, \tag{5.27}$$

where $\mathbf{e}_{kl} = [\cos\alpha_{kl} \;\; \sin\alpha_{kl}]^T$. The angles α_{kl}, $l = 1, 2, \cdots, |\mathcal{A}_k|$ determine the orientations of the localized nodes relative to the node k. The approximate CRLB can be computed using

$$\mathbf{C}_k \geq \tilde{\mathbf{I}}^{-1},$$

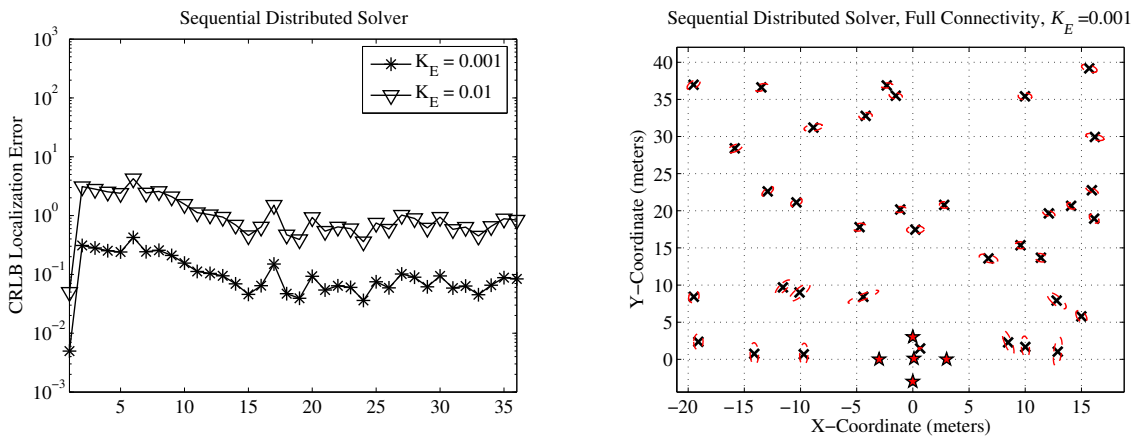
where the "effective" Fisher information matrix $\tilde{\mathbf{I}}$ is given by

$$\tilde{\mathbf{I}} = \begin{bmatrix} \sum_{l=1}^{|\mathcal{A}_k|} \frac{\cos^2(\alpha_{kl})}{\rho_{kl}^2} & \sum_{l=1}^{|\mathcal{A}_k|} \frac{\cos(\alpha_{kl})\sin(\alpha_i)}{\rho_{kl}^2} \\ \sum_{l=1}^{|\mathcal{A}_k|} \frac{\cos(\alpha_{kl})\sin(\alpha_{kl})}{\rho_{kl}^2} & \sum_{l=1}^{|\mathcal{A}_k|} \frac{\sin^2(\alpha_{kl})}{\rho_{kl}^2} \end{bmatrix}. \tag{5.28}$$

The localization error of individual nodes can be computed using the above Fisher information matrix, as in Appendix 5A.

Figure 5.8 illustrates the performance of distributed sequential location estimation for the same scenario as Figures 5.2 and 5.3. We observe that the localization error as a function of distance from the true anchors first increases and then decreases. This trend can be explained by the fact that as the distance between unlocalized nodes and the true anchors increases, each unlocalized node can receive range estimates from a larger set of localized nodes, resulting in the decrease of localization error. As in the one-dimensional case, the approximate CRLB based on the equivalent range error model— that assumes *unbiased location estimates*— predicts that the propagation of error is offset due to the availability of an increasing number of range estimates.

Figure 5.9(a) compares the three estimation approaches in terms of the average localization error as a function of the distance from true anchors, averaged over a large number of realizations of unlocalized node locations, assuming full connectivity. We see that the performance of the sequential approach is predicted to be superior to those of the fully distributed and centralized approaches, as the distance from true anchors increases. This is contrary to the expected trend, as the centralized approach should outperform the sequential approach in



(a) $N = 36$, $M = 5$

(b) $N = 36$, $M = 5$

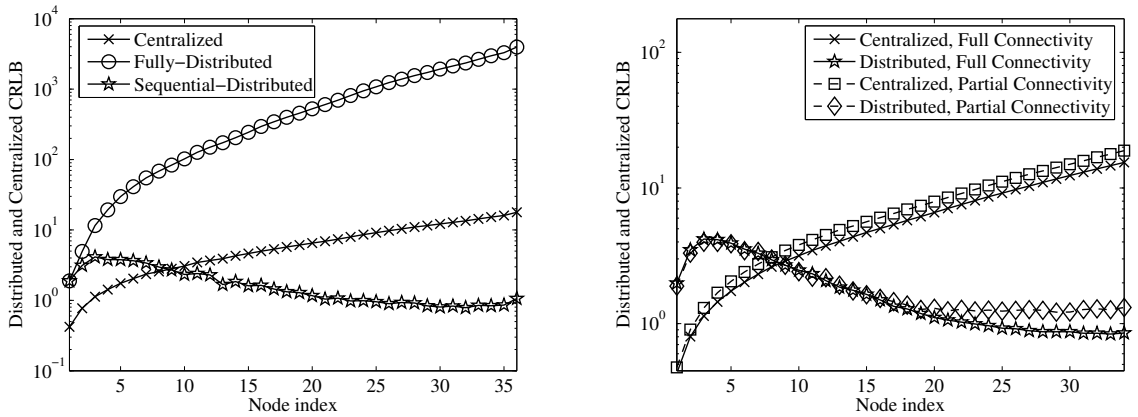
Figure 5.8: The area of interest is an $L \times L$ region, $L = 50$ meters. $N = 36$ unlocalized nodes are uniformly (randomly) distributed over this region. The unlocalized nodes, $i = 1, 2, \cdots, N$ are sorted with increasing distance from the centroid of anchors. (a) The localization error versus sorted node index, and (b) the 99% error ellipses corresponding to the sequential CRLB.

terms of average localization accuracy. As we will find in the following sections, the centralized estimation in practice does outperform the sequential approach, and this highlights the inability of the equivalent range error model to accurately describe the behavior of error propagation. As discussed in Section 5.5, the counter-intuitive behavior predicted by the equivalent range error model for anchor position errors gives us some important insights into the factors causing the propagation of localization error.

Figure 5.9(b) compares the CRLB for centralized and sequential estimation with full and partial connectivity. We see that as the extent of connectivity decreases, the localization accuracy of both methods degrades. In the following section, we compare the performance of practical centralized and sequential location estimators.

5.4 Practical Centralized and Distributed Solvers

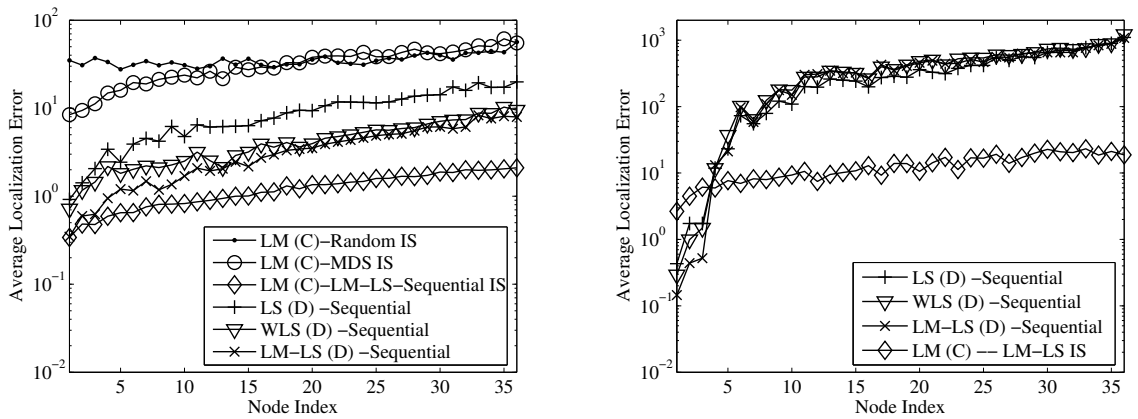
The most popular centralized location estimation approaches are based on multi-dimensional scaling (MDS) [6], [78]. However, the use of classical MDS [6] (i) requires full network connectivity and (ii) requires very accurate range information. In cases where full connectivity is not available, the sum of hop distances can be assumed [6] to be the distance between two nodes. This results in considerable degradation in localization accuracy. As a result,



(a) $N = 36$, $M = 5$, Full Connectivity

(b) $N = 36$, $M = 5$, Partial Connectivity

Figure 5.9: $L = 50$, $N = 36$. $K_E = 0.001$, Full Connectivity. The area of interest is an $L \times L$ region, $L = 50$ meters. $N = 36$ unlocalized nodes are randomly (uniformly) dispersed over the area of interest. Localization Error versus distance averaged over a large number of realizations of the locations of nodes. The value of K_E used in this simulation was 0.001. (a) Full connectivity, (b) Centralized and sequential estimators, full and partial connectivity with $R_{\max} = 30$ meters.



(a) $N = 36$, $M = 5$, Full Connectivity

(b) $N = 36$, $M = 5$, Partial Connectivity

Figure 5.10: Comparison of sequential and centralized location estimators: (a) $K_E = 0.001$, Full connectivity, (b) $K_E = 0.001$, Partial connectivity with $R_{\max} = 40$ meters. The area of interest is an $L \times L$ region, $L = 50$ meters. $N = 36$ unlocalized nodes are randomly (uniformly) dispersed over the area of interest. In the above figures, methods labeled "C" are centralized solutions, whereas methods labeled "D" are distributed methods.

MDS estimates have been used as initial solutions for iterative refinement [6] yielding good results. A brief description of the basic MDS method is provided in Appendix 5B.

We compare the performance of sequential distributed solvers based on the estimators detailed in Chapter 4, with a centralized solver based on the Levenberg-Marquardt (LM) iteration discussed in Chapter 4. The LM iterative method "LM (C)" method is applied to the following non-linear minimization problem:

$$\min \epsilon\left(\mathbf{X}\right) = \sum_{i,j} c_{ij} \left(\|\mathbf{x}_i - \mathbf{x}_j\| - r_{ij}\right)^2 . \tag{5.29}$$

where $\mathbf{X}$ is defined in (5.4), and $c_{ij} \in \{0, 1\}$ represents the absence or presence of connectivity between nodes i and j. The details of the LM iteration can be found in Section 3.6. As discussed in Chapter 4, the performance of iterative estimation schemes, in general, depends on the choice of initial solution. Figures 5.10(a) and 5.10(b) compare the performance of sequential distributed schemes with the LM (C) iteration using various initial solutions (IS).

In Figure 5.10(a), the performance of the centralized LM estimator using random or MDS-based initial estimates is very poor compared to the case with estimates from a sequential approach. It must be pointed out that MDS cannot be used unless full connectivity is assumed. We see that the performance of centralized LM method with sequential LS estimates is much better than with MDS estimates as initial solutions. However, we also note that

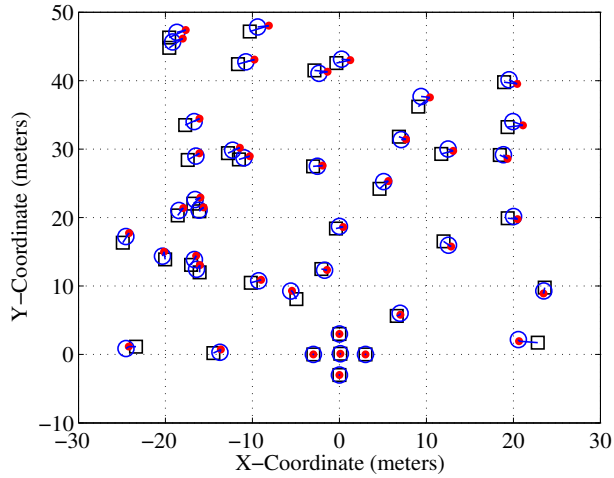
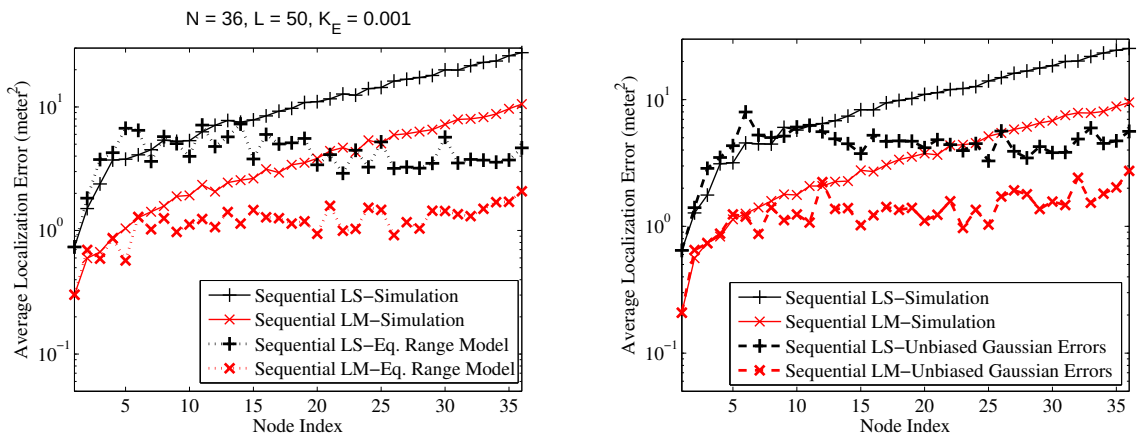


Figure 5.11: The true locations of $N = 36$ nodes are denoted by the red dots. The centralized-LM estimates are denoted by circles, and the sequential distributed LS estimates are denoted by squares. We see that the LM iteration helps drive the estimates closer to the true solutions.

the centralized scheme that uses sequential distributed estimates as the initial solution[2] provides considerable performance gains relative to the sequential distributed schemes (Figures 5.10(a) and 5.10(b)). Thus, centralized iteration using additional range estimates between unlocalized nodes can considerably improve average localization accuracy relative to distributed sequential approaches. Figure 5.11 illustrates an example of the location estimates obtained using the distributed LS and the centralized-LM iteration on the distributed LS estimates. We see that the centralized iteration drives the location-estimates closer to the true locations thereby reducing the localization error.

Figure 5.10(b) compares the performance of the distributed and centralized approaches when unlocalized nodes have connectivity with other nodes that are within a radius $R_{\max} = 40$ meters. While the performance of both estimation approaches degrades as the extent of connectivity goes down, centralized estimation is more robust to decrease in connectivity than distributed estimation. The observations made above with practical estimators are contrary to those predicted in Section 5.3, where, based on the approximate CRLB developed for anchor position errors, the sequential approach was expected to outperform the centralized approach. In the following section, we examine the accuracy of the equivalent range error model and the impact of geometry on the propagation of localization error.

[2]The sequential estimates can be computed in a centralized manner for an initial solution, prior to the LM iteration.



(a) (b)

Figure 5.12: (a) Comparison of the performance of LS and LM Sequential estimators with and without the equivalent range model, (b) Comparison of the simulated performance of LS and LM Sequential estimators with and the case with equivalent unbiased Gaussian anchor position errors. The value of $K_E = 1 \times 10^{-3}$.

5.5 CRLB, Geometry, and Design Implications

In chapter 4, the CRLB for location estimation using unbiased Gaussian range estimates and exact anchor location-estimates was derived. The trends displayed by the CRLB were seen to be similar to the trends shown by practical estimators in terms of the impact of geometry, number of range estimates, and the variances of range estimates. However, the approximate CRLB for sequential location estimation derived in Section 5.3 predicts that the increased number of range estimates from nodes that have estimated their locations causes the localization error not to diverge (Figure 5.9) as the distance from true anchors increases, while this effect was not observed with the practical estimators discussed in Section 5.4.

The derivation of the approximate CRLB used the equivalent range error model for anchor position errors due to the intractability of the *exact* CRLB that combines the effects of both range errors and anchor position errors simultaneously. One of the key assumptions in this derivation was that the anchor position errors were unbiased and Gaussian distributed. Figure 5.12(a) shows the performance of the practical sequential methods when the equivalent range model is used to replace the anchor position errors. The anchor position errors are replaced by equivalent Gaussian range errors using (5.19). We see that when this model is applied, the trends displayed by practical estimators are similar to those displayed by the approximate CRLB, and the localization error does not increase.

This suggests that the disparity between trends observed with the approximate CRLB and practical estimators arises due to the inability of the equivalent range errors to accurately model anchor position errors obtained when practical estimators are used. An explanation for this disparity is that when practical estimators are used, the coordinates of the location estimates are not necessarily unbiased or Gaussian, as seen in Figure 3.9. This explanation gains further credibility through Figure 5.12(b), where the simulated performance of the sequential LS and LM estimators is compared with the case where the actual anchor position errors are replaced by zero-mean Gaussian random variables that produce the same localization error. We see that the average localization error with the zero-mean Gaussian anchor position errors is similar to that predicted by the equivalent range error model in Figure 5.12(a), and tends to converge with increasing distance from the true anchors. Thus, a key factor that determines the propagation of localization error in practical sequential approaches is the bias in location estimates obtained through practical estimators. Therefore, a potential means of containing the propagation of localization error is to use estimators that produce location estimates that are unbiased, along with an increasing number of range estimates as the distance to the true anchors increases. Finally, we may conjecture that if the exact CRLB with both range and anchor position errors were computed, the trends would be similar to the practical estimators discussed, and would be outperformed by the centralized case.

5.5.1 Impact of geometry

It was seen in Section 5.2.2 that even when a centralized estimation approach is used, the average localization error in the ad hoc network architecture increases with distance from the true anchors (Figure 5.3(a)). This suggests that the inherent geometry of nodes in ad hoc network architecture and the propagation of localization error may be correlated. In order to demonstrate that geometry plays a significant role in dictating the propagation of localization error, consider the examples shown in Figures 5.13(a) and 5.13(b). Figure 5.13(a) presents scenario 1 (S1), where the unlocalized nodes are located at $\{(iR, 0)\}$, $i = 0, 1, 2, \cdots$, and the anchor nodes are located at $\{(-R, 0)\} \cup \{(iR, \pm R)\}$, $i = 0, 1, 2, \cdots$. The transmission radius $R_{\max}$ is selected such that the ith unlocalized node can estimate its location based on using range and location estimates from two anchors at $\{(iR, \pm R)\}$ and unlocalized node $(i-1)$. Figure 5.13(b) presents scenario 2 (S1) which is identical to scenario 1, except for an improvement in geometrical dilution of precision of anchors at each unlocalized node. The LS estimator was used to sequentially estimate the locations of unlocalized nodes, assuming unbiased Gaussian range estimates. Figure 5.13(c) shows the simulated LS localization error for the two scenarios. We see that the improvement in geometry can result in the difference between the convergence and divergence of the localization error.

The impact of geometry in the ad hoc network architecture is further demonstrated in Figure 5.14. In Figure 5.14(a), $M = 8$ anchor nodes are placed in a small region outside the area of interest. In this case, as we have seen in Figure 5.10, the use of sequential estimation

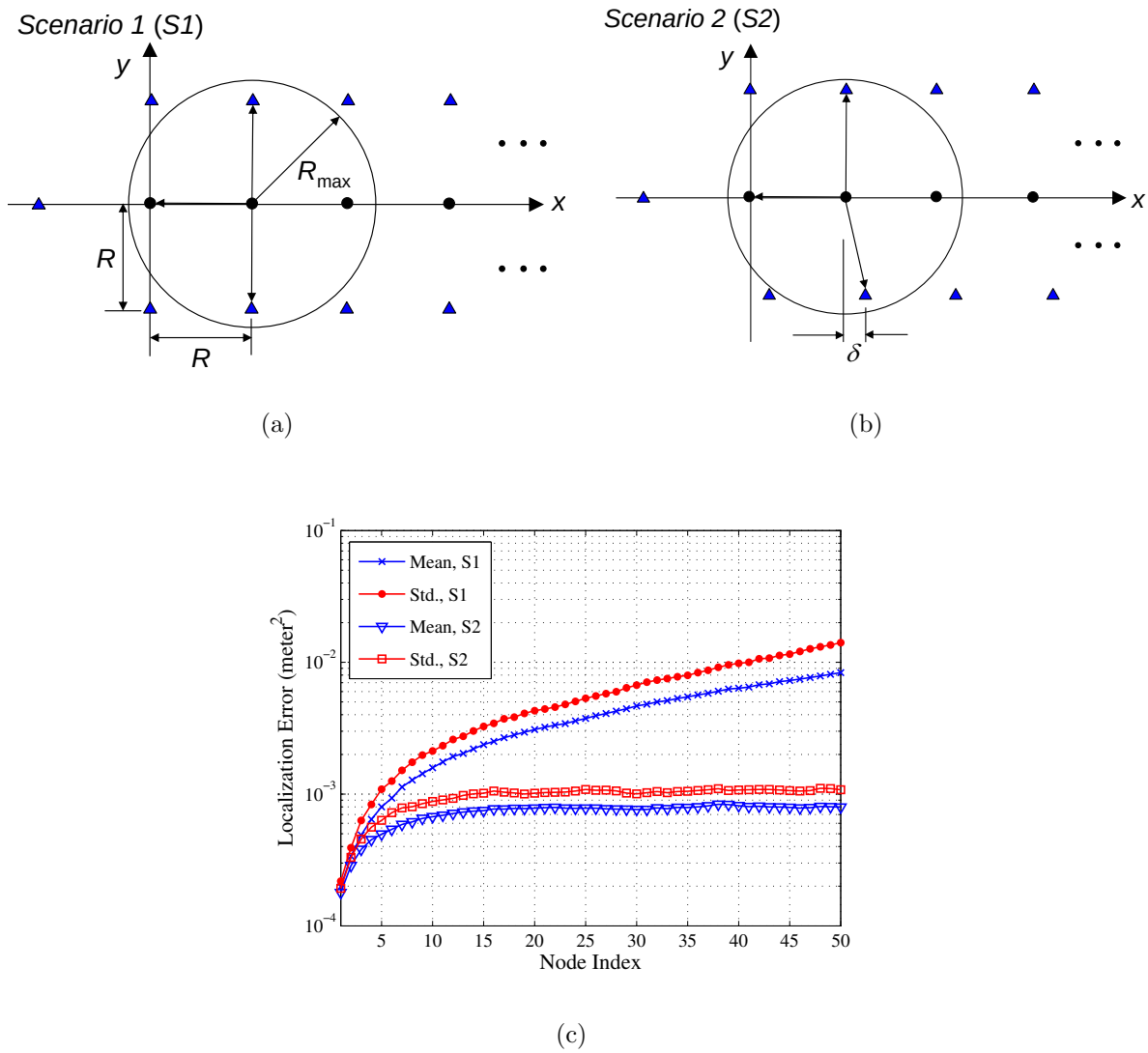


Figure 5.13: (a) Scenario 1 (S1) Here, unlocalized nodes are located on the x-axis at $x = iR$, $i \geq 0$, $R = 10$ meters, and anchor nodes are located at $(iR, \pm R)$ and $(-R, 0)$. The node transmission radius $R_{\max} = 12$ meters. (b) Scenario 2 (S2) In this case, unlocalized anchor nodes are located at (iR, R), $(iR + \delta, -R)$, $i \geq 0$ and $(-R, 0)$, $\delta = 2$ meters. (c) The mean and standard deviation of the simulated LS localization error (in meter^2) as a function of the unlocalized node index i for scenarios S1 and S2. Here, $K_E = 1 \times 10^{-4}$.

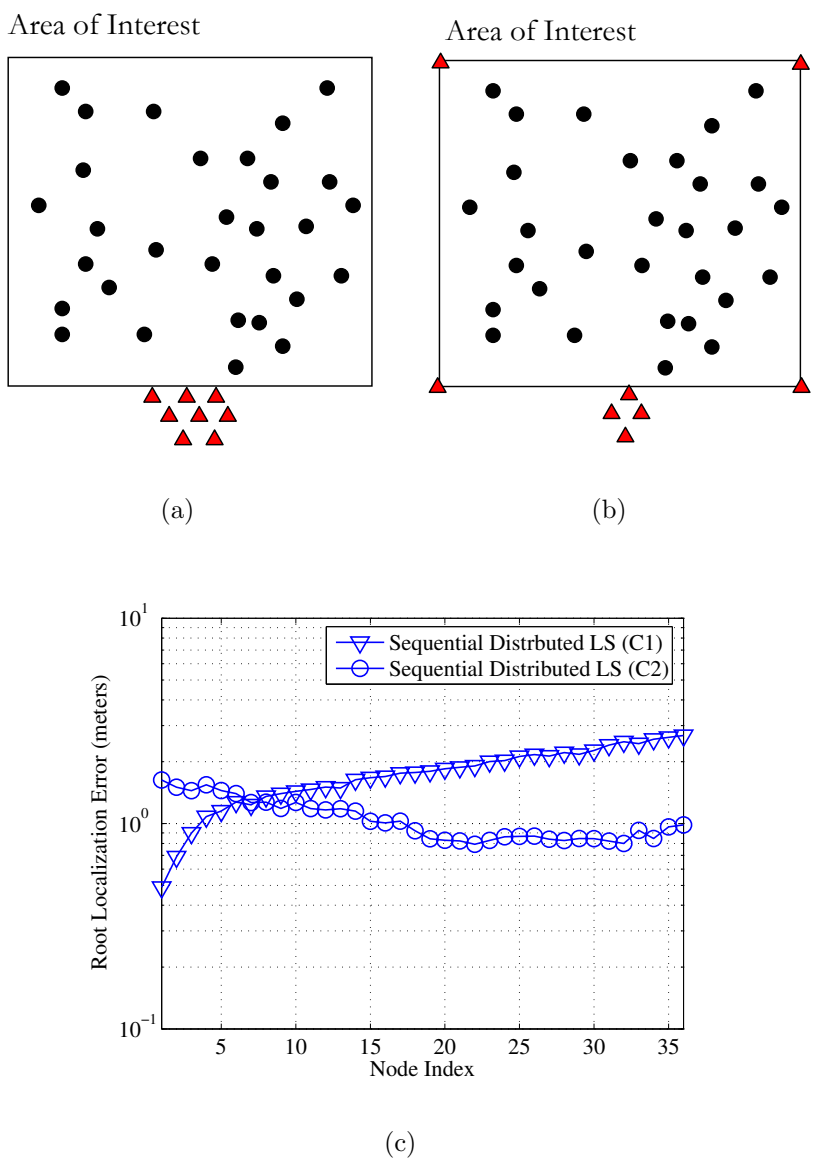


Figure 5.14: (a) Case 1 (C1) In this case, $M = 8$ anchor nodes are placed in a small area located outside the $L \times L$ area of interest, where $L = 50$ meters, (b) Case 2 (C2)4 anchor nodes of the 8 anchor nodes in C1 are moved to the corners of the area of interest, (c) the average root localization error (meters) for cases C1 and C2 using a sequential LS estimator, with $N = 36$ unlocalized nodes uniformly distributed over the area of interest, partial connectivity with $R_{\max} = 40$ meters, and $K_E = 0.001$.

approaches results in the propagation of error. However, if 4 of the 8 anchor nodes are moved to the corners of the area of interest, as in Figure 5.14(b), the resulting geometry of anchors results in lower geometric dilution of precision for unlocalized nodes within the area of interest. As seen in Figure 5.14(c), even with partial connectivity, this change in the geometry considerably mitigates the propagation of error. Thus, the propagation of error resulting from sequential location estimation can be considerably mitigated by distributing the locations of a sufficient number of anchors *around* the area of interest in a manner that ensures a low geometric dilution of precision.

5.5.2 Limits on the Area of Propagating Location Information

In the previous subsection, we saw the inherent geometry of the ad hoc network architecture results in the propagation of localization error when practical sequential estimators are applied. The most significant implication of the propagation of localization error to the design of ad hoc PoLoNets is that for a target localization accuracy, the area over which the desired localization accuracy can be guaranteed is limited. For instance, let us assume that the desired average root localization error (in meters) is specified as $\Omega_0 = 2$ meters. Further, suppose that $N = 36$ unlocalized nodes are uniformly (randomly) distributed over the area of interest, which is an $L \times L$ region, where $L = 50$ meters. Assuming $K_E = 0.001$, full connectivity, LOS links, and that the sequential LS estimator is used, then from Figure 5.15(a), the average distance $\tilde{L}$ from the centroid of true anchors at which Ω_0 can be maintained is approximately $\tilde{L}_{LS} = 20$ meters. If the dimension of the area of interest L is larger than $\tilde{L}$, then the average localization error in some regions of the area of interest will be larger than desired.

As is evident from Figure 5.15(a), two ways of expanding the area over which the desired localization accuracy can be guaranteed are (i) using a superior location estimator, (ii) increasing node density, and/or (iii) increasing transmit power. We see that if the sequential LS estimator is replaced by the sequential WLS estimator, while fixing N and the other parameters, the value of $\tilde{L}_{WLS} \approx 40$ meters. Increasing N from 36 to 72 while keeping L fixed further improves performance: $\tilde{L}_{LS} \approx 30$ meters and $\tilde{L}_{WLS} \approx 60$ meters.

In Figure 5.15(b), we see that connectivity plays an important role in determining $\tilde{L}$. We see that if $R_{\max}$ denotes the transmission radius, the average localization error increases sharply after a distance $\sim R_{\max}$ from the true anchors. For this value of $R_{\max}$, the value of $\tilde{L}_{WLS}$ decreases from 40 meters to 30 meters when $N = 36$, and from 60 meters to 35 meters when $N = 72$.

Based on the observations in Figures 5.15(a) and 5.15(b), we can derive linear models for the propagation of localization error, and its impact of the area over which a desired localization accuracy can be ensured. The log localization accuracy (in terms of the localization error or root localization error) can be modeled as linear function of $\tilde{L}$ as shown in Figure 5.15(c). From Figure 5.15(a), we see that the slope of the log localization accuracy is approximately

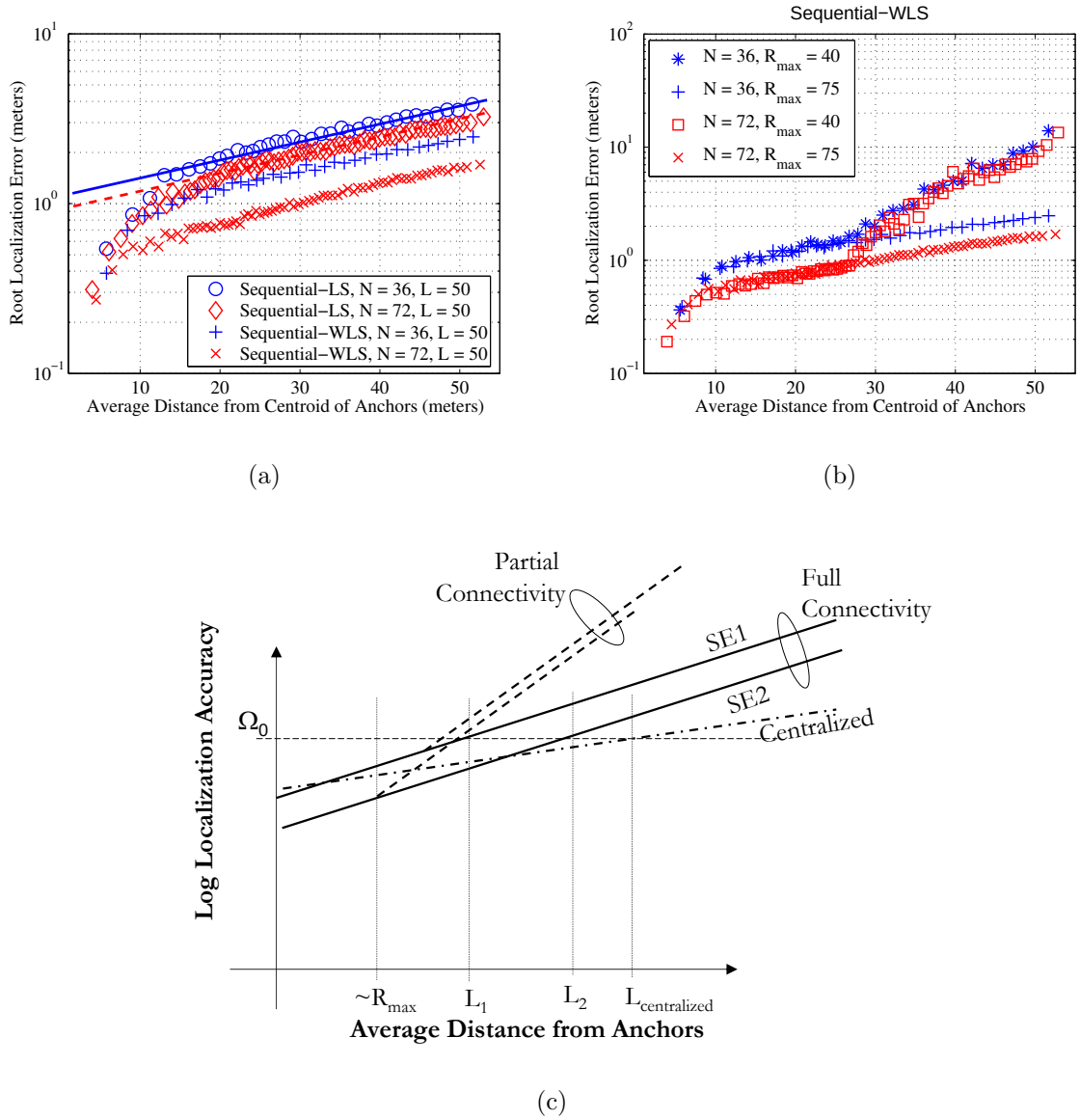


Figure 5.15: Modeling of the propagation of localization error in ad hoc scenarios with practical sequential estimators: (a) sequential LS and WLS estimators with full connectivity and different node densities, (b) sequential WLS estimator with partial connectivity and different node densities, (c) Linear models for log localization accuracy.

the same across the discussed sequential estimators. In the absence of full connectivity, we see from Figure 5.15(b) that the slope of the log localization accuracy increases when $\tilde{L} \lesssim R_{\max}$. For such cases, a two-slope linear model for the log localization accuracy can be used as shown in Figure 5.15(c).

From the above observations we can enumerate various ways of increasing the area over which the desired localization accuracy can be guaranteed when sequential estimation is used: (i) improving range accuracy, (ii) using a superior (and possibly more complex) location estimator, (iii) increasing node density, (iv) increasing transmit power, (v) improving the geometry of anchor nodes, (vi) using methods to mitigate the propagation of error (discussed in the following section). Using the listed methods, if the desired localization accuracy still cannot be guaranteed over the entire region of interest, then switching to a centralized estimation approach may be considered as shown in Figure 5.15(c).

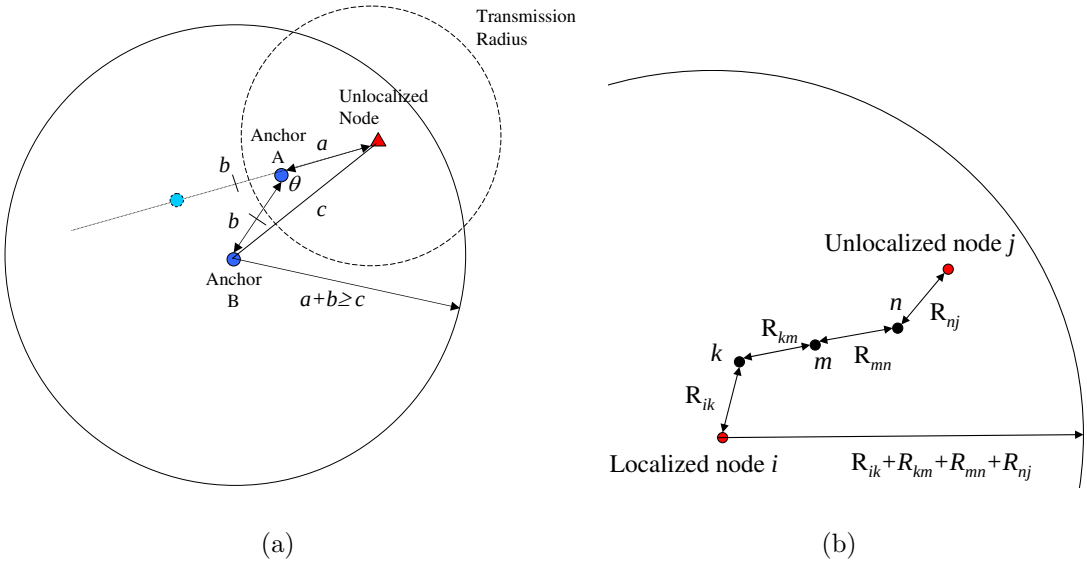
In the discussion of the propagation of localization error thus far, the links between nodes were assumed to be LOS in nature. The propagation of localization error is likely to exacerbate in the presence of NLOS links. In the following section, we discuss a novel method of mitigating the propagation of localization error, especially in NLOS propagation environments.

5.6 Multi-hop range-based propagation of error mitigation

The propagation of error problem and its mitigation are discussed in [75], and briefly in [76]. In [75], radio implementation constraints on the maximum angle-of-arrival errors and range measurement errors from one-hop neighbors of a given node are used to constrain the set of possible locations for the node. In [76] the covariance matrix of the anchor position errors is relayed from node to node to provide a better estimate of the weighting matrix in the WLS formulation.

In the previous section, we saw that the extent of connectivity plays an important role in the propagation of error. We observed that with both centralized and sequential modes of estimation, as the extent of connectivity decreases, the localization error increases. Further, the range estimates were assumed to unbiased, implying LOS links between nodes. In NLOS environments, the propagation of error can be considerably exacerbated due to the biased nature of NLOS range estimates.

We found previously that the biased nature of location estimators contributed significantly to the propagation of error. As shown in Chapter 4, the effects of bias can be limited by constraining the feasible region for location-estimates. In this section, we propose such a means of partially mitigating the propagation of error, that incorporates (i) the possible NLOS nature of range estimates, and (ii) partial connectivity. This method uses NLOS



(a) (b)

Figure 5.16: (a) Illustration of the use of two-hop distances to provide bounds on node distance, (b) Illustration of bounds in the multi-hop case.

range estimates and *multi-hop* distances between two nodes to bound the possible set of node locations. No angle-of-arrival information is assumed. Consider the scenario illustrated in Figure 5.16(a). Here, the unlocalized node can obtain range estimates from its one-hop localized neighbor A, but not from its two-hop localized neighbor B. However, A and B are one-hop neighbors, and A can estimate the distance between itself and B. We shall demonstrate that the range estimates from two-hop localized neighbor B can be used to constrain the location of the unlocalized node.

In Figure 5.16(a), if $\|\mathbf{x} - \mathbf{x}_A\| = a$, $\|\mathbf{x}_A - \mathbf{x}_B\| = b$, and $\|\mathbf{x} - \mathbf{x}_B\| = c$, we see that

$$c^2 = a^2 + b^2 - 2ab\cos\theta.$$

From the above equation (or the Triangle inequality),

$$c \leq a + b,$$

with equality achieved when $\theta = \pi$. This implies that when anchor nodes A and B are collinear, the sum of the distances a and b is also equal to the distance between anchor node B and the unlocalized node. Otherwise, the sum of the distances a and b is always larger than the true distance between node B and the unlocalized node. This implies that we can bound the location of the unlocalized node using the sum of single-hop distances from an anchor node to the unlocalized node. We can ensure that the unlocalized node lies *within* this distance between using a multiplicative factor $\alpha > 1$:

$$c < \alpha(a + b).$$

The factor α also makes the bounds robust in the presence of range measurement noise and anchor position errors. The above upper bound becomes tighter as the angle θ increases and tends toward π. The above result can further be extended to the general multiple two-hop localized anchors case as follows: If $\mathbf{x}_j$ is the location of an anchor, then the location of an unlocalized node at $\mathbf{x}_i$, and $c_{ij} = 0$ (i.e., node i and node j are not one-hop neighbors), then

$$\|\mathbf{x}_i - \mathbf{x}_j\| \leq R_{ik} + R_{kj},\ c_{ij} = 0,\ c_{ik} = 1,\ c_{kj} = 1, \tag{5.30}$$

where c_{mn} are elements of the network connectivity matrix. This can also be extended to the three-hop case:

$$\begin{aligned} \|\mathbf{x}_i - \mathbf{x}_j\| &\leq R_{ik} + R_{kl} + R_{lj},\ c_{ij} = 0,\ c_{ik} = 1,\ c_{kl} = 1,\ c_{lj=1}, \\ \Rightarrow \|\mathbf{x}_i - \mathbf{x}_j\| &< \alpha\left(R_{ik} + R_{kl} + R_{lj}\right),\ c_{ij} = 0,\ c_{ik} = 1,\ c_{kl} = 1,\ c_{lj=1},\ \alpha > 1. \end{aligned} \tag{5.31}$$

The above bound is illustrated in Figure 5.16(b). Each of the constraints in (5.31) can be used to create a feasible region for the estimate of the unlocalized node in order to limit the propagation of error. The number of constraints can be increased by using distance constraints over a larger number of hops, which further reduces the size of the feasible region. The gains achieved by increasing the number of constraints come at the price of computational complexity. Figure 5.17 compares the performance of the sequential LS estimator with the sequential LP approach assuming LOS links, where $N = 36$, $M = 5$ and $L = 50$ meters. We see that the LP approach can provide gains in localization accuracy: the distance from the true increases at which an average root localization error $\Omega_0 = 2$ meters can be provided increases from $\tilde{L}_{LS} = 20$ meters to $\tilde{L}_{LP} = 30$ meters.

However, the main advantage of the discussed sequential LP scheme is in NLOS environments. We see that the constraints in (5.31) are similar to the NLOS constraints discussed in Chapter 4. This implies that the above constraints can be linearized as seen in Section 4.7.2, and combined with the NLOS constraints in the linear-program formulated in Sections 4.7.3, constraining the feasible region for a given node's location. The increase in computational complexity is due to the increase in the number of constraints. This results in the straightforward formulation of a *sequential*-LP scheme that uses LOS range information from available localized nodes to define the objective function and constraints, with NLOS and multi-hop constraints defining the feasible region.

Figure 5.18 compares the performance of the sequential distributed LS method with the sequential-LP method that uses two-hop constraints on a node's location. The simulation parameters are the same as those used in Figure 5.10, with partial connectivity, $R_{\max} = 30$ meters, and the probability that a given link is NLOS is p. Only two-hop constraints were considered in the formulation of the sequential-LP, and the sequential-LS estimator uses only LOS range estimates. We see that as p increases, the localization accuracy degrades, and the advantage of the sequential-LP scheme over the sequential-LS scheme progressively increases. When $p = 0.1$, the average distance from the true anchors at which a root localization error $\Omega_0 = 2$ meters can be achieved, increases from $\tilde{L}_{LS} = 15$ meters to $\tilde{L}_{LP} = 25$ meters. The

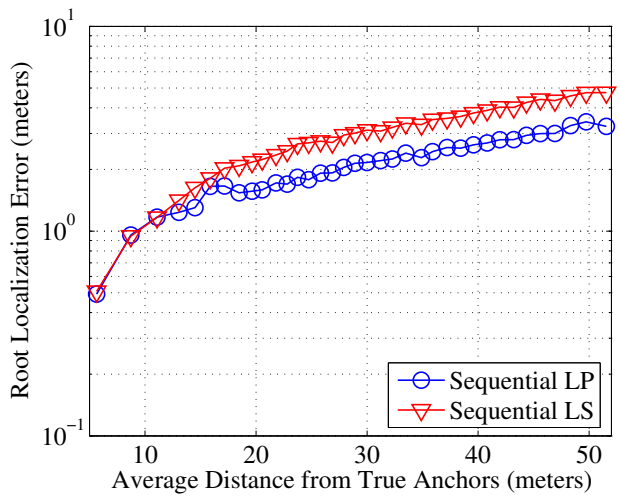


Figure 5.17: Gains in the average root localization error when the LP-based mitigation scheme is applied with all LOS range estimates. Here, $N = 36$, $M = 5$, $L = 50$ meters, $K_E = 0.001$. Full connectivity is assumed.

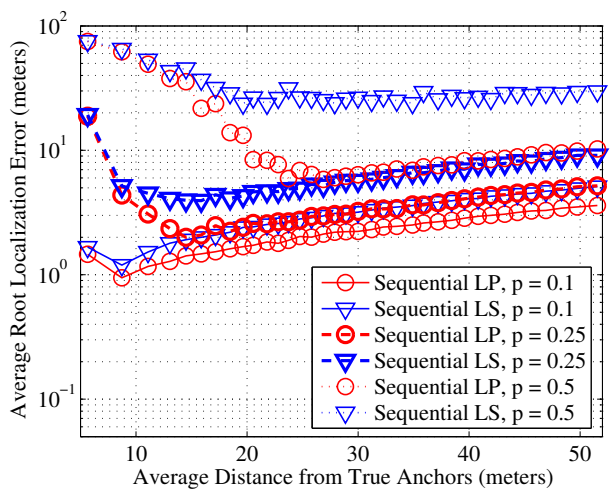


Figure 5.18: Comparison of the performance of the sequential-LP and sequential-LS approaches with NLOS links (with probability p) and partial connectivity (with $R_{max} = 50$ meters. As before, $N = 36$, $M = 5$, $K_E = 0.001$.

error for a small number of nodes is large, as the links to the true anchors may be NLOS, and only a small number of LOS range estimates may be available. Thus, in order to prevent large localization error, LOS links between the true anchors and nearby reference nodes need to ensured.

5.7 Conclusions

In this chapter, we presented an in-depth study of the propagation of error problem when sequential estimation, which achieves a trade-off between localization accuracy and communication cost vis-a-vis the fully distributed and centralized approaches, is used in an ad hoc network architecture. The propagation of error in sequential location-estimation restricts the area of the region over which a desired average localization accuracy can be achieved. It was found that the geometry of anchors and the bias resulting from the use of practical estimators were important factors contributing the propagation of error. We discussed the various ways of increasing the area over which the desired localization accuracy can be guaranteed when sequential estimation is used: (i) improving range measurement accuracy, (ii) using a superior (with minimal bias) location estimator, (iii) increasing node density, (iv) increasing transmit power, (v) improving the geometry of anchor nodes, (vi) using methods to mitigate the propagation of error. A novel method of mitigating the propagation of localization error based on linear-programming that incorporates NLOS range estimates was proposed. In the following chapter, we present measurement results that evaluate the performance of the proposed method in mitigating the propagation of error.

5.8 Appendix 5A: CRLB for Centralized Location-Estimation

The vector of parameters to be estimated is

$$\mathbf{X} = \begin{bmatrix} x_1 & y_1 & x_2 & y_2 & \cdots & x_N & y_N \end{bmatrix}^T_{2N\times 1}.$$

The given set of range estimates $\mathcal{R}$ can be cast into the vector γ. As each range estimate in γ is independent, we can write the conditional p.d.f. of γ as:

$$\begin{aligned}
f(\gamma|\mathbf{X}) &= \prod_{i,j} \frac{1}{\sqrt{2\pi\sigma_{ij}^2}} \exp\left(-\frac{(R_{ij}-\gamma_{ij})^2}{2\sigma_{ij}^2}\right) \\
\Rightarrow \log f(\gamma|\mathbf{X}) &= \log\left(\prod_{i,j} \frac{1}{\sqrt{2\pi\sigma_{ij}^2}}\right) - \sum_{i,j} \frac{(R_{ij}-\gamma_{ij})^2}{2\sigma_{ij}^2} \\
\Rightarrow \log f(\gamma|\mathbf{X}) &= \log\left(\prod_{i,j} \frac{1}{\sqrt{2\pi\sigma_{ij}^2}}\right) - \sum_{i,j} \frac{\left(\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}-\gamma_{ij}\right)^2}{2\sigma_{ij}^2},
\end{aligned}$$

where γ_{ij} is the available range estimate between nodes i and j. The entries of the $(2N\times 2N)$ Fisher information matrix are given by

$$\mathbf{I} = -E_\gamma\left\{\frac{\partial^2 \log f(\gamma|\mathbf{X})}{\partial \mathbf{X}^2}\right\}$$

It has been shown [19] that

$$\begin{aligned}
I_{2i-1,2j-1} &= \begin{cases} \sum_{j\in\mathcal{T}_i} \frac{\cos^2\alpha_{ij}}{\sigma_{ij}^2} & i=j, \\ -\frac{\cos^2\alpha_{ij}}{\sigma_{ij}^2} & i\neq j,\ j\in\mathcal{U}_i, \\ 0 & \text{otherwise.} \end{cases} \\
I_{2i,2j} &= \begin{cases} \sum_{j\in\mathcal{T}_i} \frac{\sin^2\alpha_{ij}}{\sigma_{ij}^2} & i=j, \\ -\frac{\sin^2\alpha_{ij}}{\sigma_{ij}^2} & i\neq j,\ j\in\mathcal{U}_i, \\ 0 & \text{otherwise.} \end{cases} \\
I_{2i-1,2j} &= \begin{cases} \sum_{j\in\mathcal{T}_i} \frac{\cos\alpha_{ij}\sin\alpha_{ij}}{\sigma_{ij}^2} & i=j, \\ -\frac{\sin\alpha_{ij}\cos\alpha_{ij}}{\sigma_{ij}^2} & i\neq j,\ j\in\mathcal{U}_i, \\ 0 & \text{otherwise,} \end{cases} \\
&= I_{2i,2j-1} = I_{2j-1,2i} = I_{2j,2i-1}. \qquad (5.32)
\end{aligned}$$

From the above result, it is straightforward to show that the total localization error is given by:

$$\Omega_t = \sum_{i=1}^{N} E\left\{\|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2\right\} = \mathrm{Tr}\left(\mathbf{I}^{-1}\right)$$

The CRLB for the location of the kth unlocalized node is given by

$$\mathbf{C}_{\mathbf{x}_l} \geq \begin{bmatrix} J_{l-1,\ l-1} & J_{l-1,\ l} \\ J_{l,\ l-1} & J_{l,\ l} \end{bmatrix}, \tag{5.33}$$

where $\mathbf{J} = \mathbf{I}^{-1}$.

5.9 Appendix 5B: Multi-Dimensional Scaling (MDS)

Multi-Dimensional Scaling (MDS) is used to provide a visualization of entities, given the distance metric for every pair of entities. Given N nodes, with the $N \times N$ distance matrix $\mathbf{D} = \{\delta_{ij}\}$, MDS attempts to estimate $\mathbf{x}_i$ that minimizes a "stress function", which is defined as:

$$S = \sum_{i,j} \left[\delta_{ij}^2 - (\mathbf{x}_i - \mathbf{x}_j)^T (\mathbf{x}_i - \mathbf{x}_j)\right]. \tag{5.34}$$

It is useful to make a few observations based on the above definition: (i) the optimization is based on the squares of the distance measures δ_{ij} and not the distance measures themselves, (ii) all distance measures are weighted equally, (iii) the distance measure between every pair of entities is required. A weighted form of MDS requires the use of majorizing functions, such as the one formulated in [78].

The basic MDS method uses the matrix of squared distance measures $\mathbf{D}^{(2)} = \{\delta_{ij}^2\}$, and performs a process called "double-centering" to obtain the matrix $\mathbf{B} = \mathbf{X}\mathbf{X}^T$, where the $N \times N$ matrix $\mathbf{X}$ is an N dimensional representation of the N entities:

$$\mathbf{B} = -\frac{1}{2}\mathbf{J}\mathbf{D}^{(2)}\mathbf{J},\ \mathbf{J} = \mathbf{I}_{N\times N} - \frac{1}{N}\mathbf{1}^T\mathbf{1}. \tag{5.35}$$

Once $\mathbf{B}$ is obtained, the matrix $\mathbf{X}$ can be extracted using singular value decomposition (or eigenvalue decomposition):

$$\mathbf{B} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T = \mathbf{X}\mathbf{X}^T \Rightarrow \mathbf{X} = \mathbf{U}\mathbf{\Lambda}^{\frac{1}{2}}. \tag{5.36}$$

The columns of $\mathbf{X}$ (sorted based on decreasing magnitude of eigenvalues) beyond two (or three) dimensions are discarded, allowing $\mathbf{X}$ to be visualized.

Chapter 6

Measurement Results

6.1 Introduction

In chapter 4, we proposed and analyzed NLOS identification and mitigation methods, while in chapter 5, we studied the effects of the propagation of localization error and techniques for its mitigation. The efficacy of the proposed methods was demonstrated through simulations with the range error models discussed in Chapter 2. In this chapter, we present indoor measurement results that (a) demonstrate that UWB PoLoNets have the potential to achieve practical target localization accuracies (~ 1 meter^2 for the target applications discussed in Chapter 1), and (b) evaluate the efficacy of the proposed NLOS identification, NLOS mitigation, propagation of error mitigation algorithms, and the use of Kalman smoothing techniques in practical scenarios.

This chapter is organized as follows: in Section 6.2, we present details of the setup employed for recording measurements that are used to evaluate the performance of the proposed algorithms. The evaluation of the NLOS identification and mitigation algorithms, as well as propagation of error mitigation methods are discussed in Section 6.3. The key findings in this chapter are summarized in Section 6.4.

Previous Work included in this chapter: Kalman smoothing for location estimates in cellular networks has been studied in [29].

Original Contributions in this chapter:

- A complete evaluation of the performance of an indoor UWB PoLoNet based on measurements.
- Validation of the proposed NLOS identification, NLOS mitigation, and propagation of error mitigation algorithms through measurement results [26].

6.2 Measurement Setup

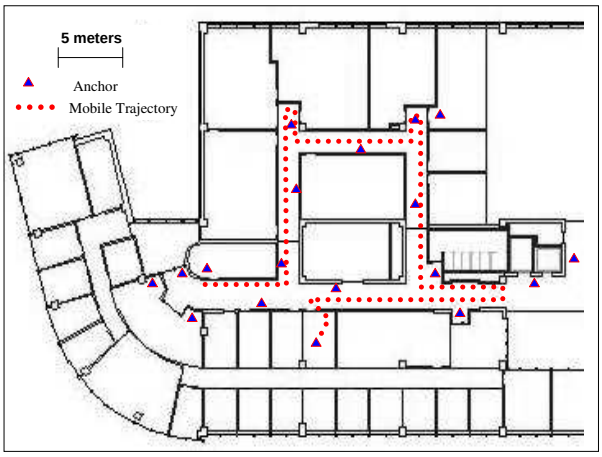
The schematic in Figure 6.1(a) shows the floor plan and the measurement locations, which comprise $N = 71$ measurement points in the mobile's trajectory, and $N_A = 18$ anchor locations as shown in Figure 6.1(c). These measurements were conducted in the corridors of the 4th floor of Durham Hall, Virginia Tech. The measurement locations in the mobile's trajectory were separated by approximately 1.2 meters. The received signal at each of these points from the transmitter placed at different anchor locations was measured, provided the signal could be captured. A Bicone antenna (see Figure 6.1(b)) connected to a 30 ps pulse generator was used as the transmitter, and the receiver consisted of a second Bicone antenna connected to a digital sampling oscilloscope that was triggered by the pulser. The pulse repetition frequency was set at 200 KHz, and the effective sampling frequency of the oscilloscope at 20 GHz. The averaged received signals from the oscilloscope were extracted, and subsequently bandpass-filtered with lower and upper cutoffs of 3.1 GHz and 10.6 GHz respectively. For each measurement location, the state of the channel (LOS or NLOS) was noted. The "soft NLOS" cases, where the LOS path was present albeit attenuated, were classified as LOS scenarios (cf. Section 2.2.4), as the TOA of the first path can still be estimated.

The distance between a given point in the mobile's trajectory and a given anchor was estimated using the corresponding received signal, if available. The range estimation algorithm was based on energy thresholding discussed in Section 2.4, with a threshold $\alpha = 36$. The range estimation method was calibrated using a reference measurement taken at a distance $d_0 = 1$ meter. For each mobile measurement location, based on the range estimates obtained from the set of available anchors, the location can be estimated using a variety of methods. These location estimates can then be compared to the physically measured location of the receive antenna. The errors arising from the physical measurement and placement of the Bicone antennas are expected to be limited to within ± 5 centimeters.

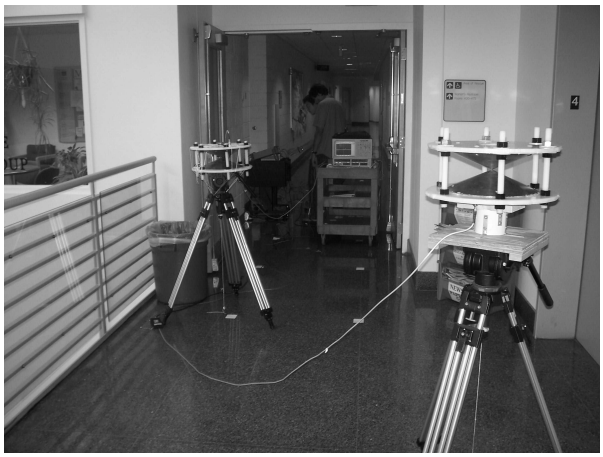
Due to the nature of measurement locations, the total number of LOS range estimates was approximately three times the total number of NLOS range estimates, i.e., $p_L = 0.75$, $p_N = 0.25$. It must be pointed out that the probability of NLOS links between nodes can be significantly higher in indoor scenarios. In the following subsections, we demonstrate the efficacy of the NLOS identification, NLOS and propagation of error mitigation schemes through measurement results.

6.3 Measurement Results

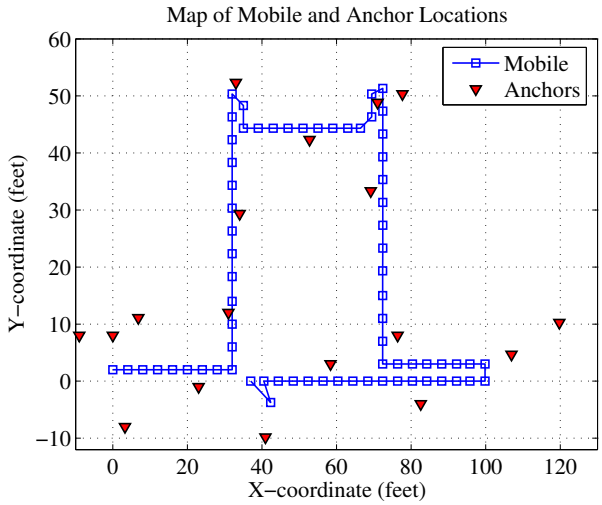
In this section, we present an evaluation of the efficacy of NLOS identification, NLOS mitigation, and propagation of error mitigation methods. In order to demonstrate the utility of the proposed algorithms, and the impact of NLOS range information in particular, we compare



(a)



(b)



(c)

Figure 6.1: (a) Floor-plan [1] of 4th floor, Durham Hall, Virginia Tech, (b) photograph of measurement setup comprising Bicone antennas, oscilloscope and pulser, and (c) the relative placement of anchor and mobile locations. The axes in (c) correspond to the x and y coordinates in feet.

the following localization methods: (a) the least-squares (LS) estimator discussed in Chapter 3, retaining only LOS range estimates and discarding NLOS range estimates after channel state identification when more than two LOS range estimates are available ("LS-(Pure LOS)"), (b) the LS estimator incorporating all available range estimates without distinguishing between LOS and NLOS range estimates ("LS-(LOS+NLOS)") and (c) the linear-programming (LP) approaches discussed in Chapter 4, that treat LOS and NLOS range estimates differently following channel state identification ("LP-basic" and "LP-extended"). The LS-(Pure LOS) method is identical to the LS-(LOS+NLOS) method when fewer than three LOS range estimates are available.

As discussed in Chapter 4, the LS-(Pure LOS) method essentially represents a conservative location-estimation method, since NLOS range estimates are discarded as they are known to be biased. The LS-(LOS+NLOS) method represents a "blind" approach, as all range estimates, unbiased or biased, are directly incorporated into the least-squares formulation. The LP-based methods utilize LOS range estimates in a way similar to the LS-(Pure LOS) method, but utilize NLOS range estimates to create a feasible region for potential solutions [27, 28], thereby resulting in the "soft-activation" of NLOS range information. These localization methods are compared in terms of the "root localization error" (in meters), defined as the Euclidean distance between the true and estimated locations of the mobile. For all three cases, if an estimate cannot be computed due to insufficient range estimates, the location-estimate is set to the previous location-estimate.

We begin by evaluating the performance of the LS and LP-based estimators when "perfect" (recorded) channel state information is available. We then apply the NLOS identification methods discussed in Chapter 4 to examine the efficacy of the proposed methods.

6.3.1 Localization Accuracy using Perfect Channel State Information

Figure 6.2 compares the location estimates of the LS and LP approaches with the physically measured node locations. The estimates obtained using the LS-(Pure LOS), LS-(LOS+NLOS), LP-basic and LP-extended methods are respectively shown in Figures 6.2(a), 6.2(b), 6.2(c), and 6.2(d) respectively. The LP-based methods clearly outperform the LS-based methods, and the trends predicted by (4.36) are observed.

Figure 6.3 compares the performance of the LS and LP-based methods in terms of the root localization error (meters) versus the index of the mobile's measurement location. The time-averaged root localization error values are also shown. The time-averaged root localization error obtained using the LP-extended method is approximately 0.49 meters. The relative gains of the LP-extended method over the LS-(Pure LOS), LS-(LOS+NLOS), and LP-basic methods in terms of the time-averaged root localization error are 282% 444%, and 6.15% respectively.

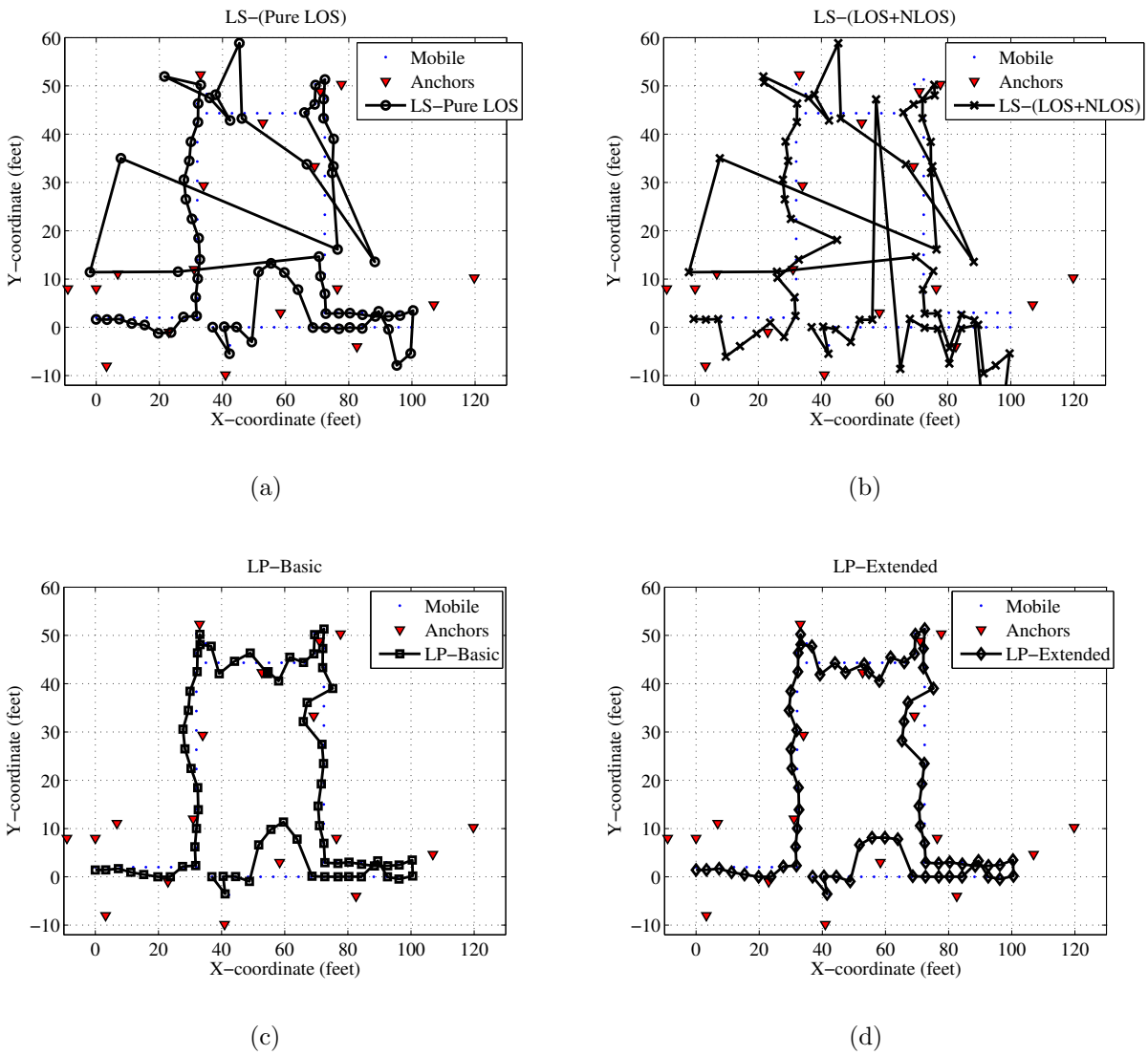
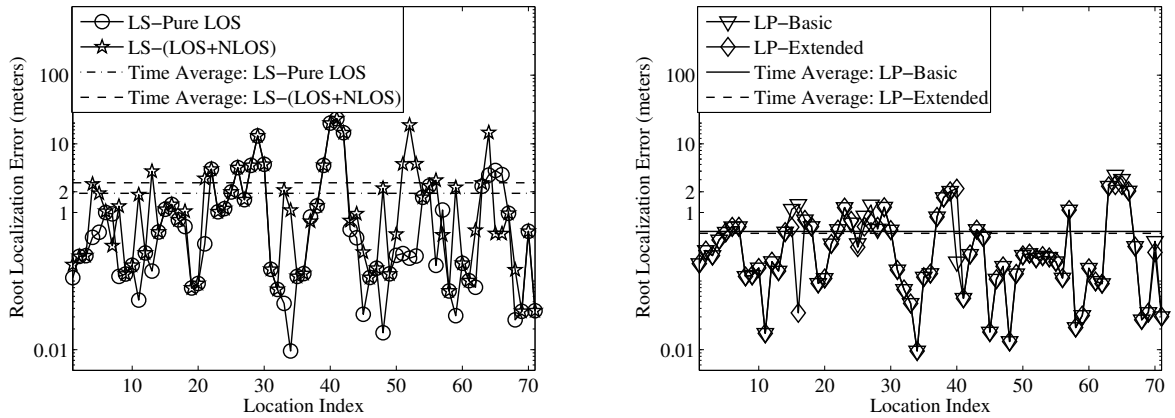


Figure 6.2: Physically measured mobile locations, and location estimates obtained using (a) LS-(Pure LOS) Estimator, (b) LS-(LOS+NLOS) Estimator, (c) LP-Basic Estimator, (d) LP-Extended Estimator. The anchor locations are also shown in each of the figures.



(a) (b)

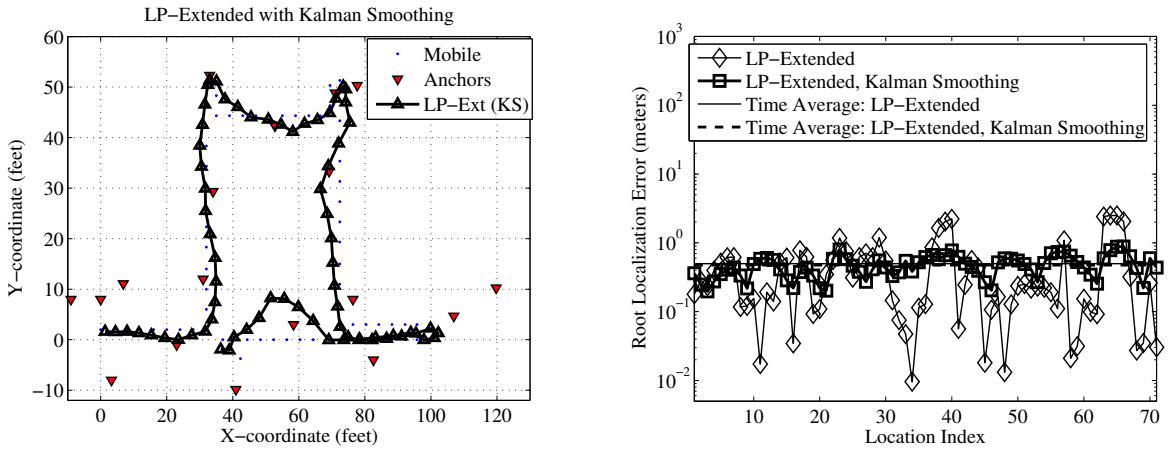
Figure 6.3: The root localization error (in meters) obtained using (a) LS-based methods, and (b) LP-based methods. The time-averaged root localization error is also shown for all cases.

From Figure 6.3(b), we see that there is considerable fluctuation in the root localization error versus the mobile's measurement location. This fluctuation can be reduced by Kalman smoothing, which may also help improve the localization accuracy in some scenarios. Kalman smoothing applied to mobile location estimation relies on a motion model and previous location estimates to reduce large-scale variations in current mobile location estimates. Kalman smoothing has previously been applied to cellular location tracking [29] and UWB localization systems [79]. The detailed derivation of the Kalman smoothing method that can be applied to mobile location estimation is presented in Appendix 6A.

Figure 6.4 compares the root localization error with and without Kalman smoothing applied to LP-extended location estimates. We see that although Kalman smoothing improves localization accuracy in some mobile locations, there is partial degradation in the localization accuracy at other locations. The gain of applying Kalman smoothing in terms of the time-averaged root localization error is 4.5%. As we shall discuss in Chapter 9, power-control schemes can also be used to reduce the fluctuation in localization accuracy.

6.3.2 Localization Accuracy with Channel Identification

The channel identification methods discussed in Chapter 4 are based on received signal statistics such as the RMS delay spread and the number of significant multipath components. Figure 6.5 shows the empirical histograms of the RMS delay spread and the number of



(a) (b)

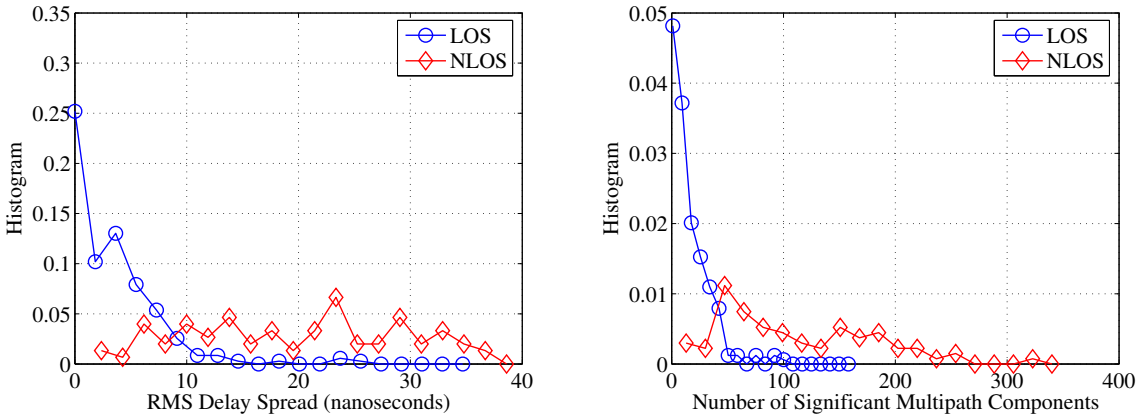
Figure 6.4: Impact of Kalman smoothing: (a) Kalman-smoothed LP-extended location estimates (cf. Figure 6.2(d)), (b) comparison of the root localization error (in meters) of LP-Extended location estimates with and without Kalman smoothing.

significant[1] multipath components corresponding to LOS and NLOS range measurements. As expected, we see that the RMS delay spread and the number of significant multipath components for NLOS propagation scenarios are, with high probability, much larger than the corresponding LOS scenarios.

In the following, we compare the performance of the LS-(Pure LOS), LS-(LOS+NLOS) and LP-extended estimators when channel state estimation through the hypothesis-testing of the RMS delay spread is applied. A threshold $T = 9.5$ nanoseconds was used for hypothesis testing the RMS delay spread. As before, we see that the LP-extended method outperforms both the LS-(Pure LOS) and LS-(LOS+NLOS) methods in terms of localization accuracy. Figures 6.6(a) and 6.6(b) compare the number of LOS and NLOS estimates respectively observed during measurement and estimated using the RMS delay spread. The total probability of error P_e in channel state estimation was found to be approximately 9% using the RMS delay spread estimate. It was verified that P_e drops to approximately 5% when *both* the RMS delay spread and the number of significant multipath components (with a threshold $T_p = 50$) are used for joint hypothesis testing.

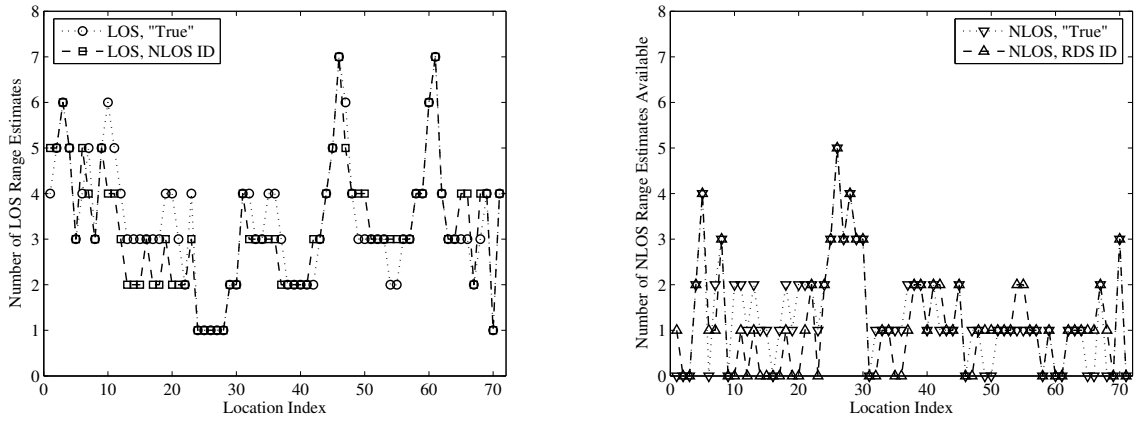
Figures 6.7(a) and 6.7(b) compare the root localization error (in meters) achieved by the three methods, with (a) "perfect" channel state information, and (b) channel state estimates based on hypothesis-testing of the RMS delay spread. We see that the use of channel state estimates based on RMS delay spread results in negligible degradation in the localization

[1]The amplitude threshold employed here was 20 dB.



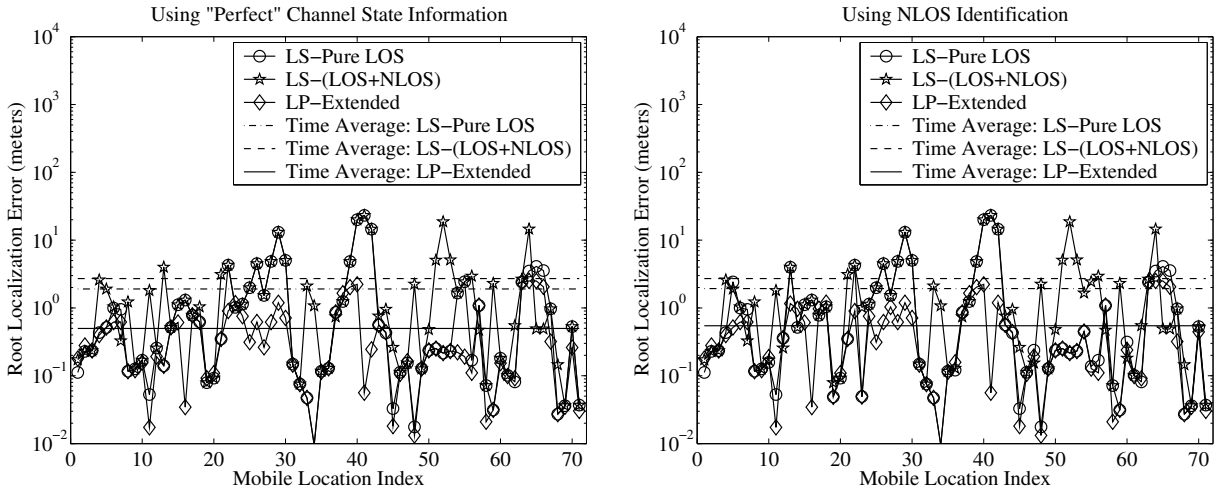
(a) (b)

Figure 6.5: Measured delay spread statistics: histograms of (a) the RMS delay spread, and (b) a) the number of significant multipath components, corresponding to LOS and NLOS scenarios.



(a) (b)

Figure 6.6: Number of (a) LOS and (b) NLOS range estimates for different mobile location observed using measurements and through channel state estimation using RMS delay spread hypothesis-testing.



(a) (b)

Figure 6.7: Root localization error (meters) versus mobile location index obtained using different methods when (a) perfect channel state information is available, and (b) when the channel state information is estimated from the RMS delay spread.

error, when compared to the case with perfect channel state information. For instance, the performance degradation in terms of the time-averaged root localization error for the LS-(Pure LOS) and LP-extended methods are 1.4% and 5.7% respectively. We further note that after channel state identification, the time-averaged localization error for the LS-(LOS+NLOS), LS-(Pure LOS) and LP-extended methods are approximately 2.71, 1.90 and 0.50 meters respectively. This suggests that while channel state identification followed by the discarding of NLOS range estimates improves localization accuracy when the LS estimator is used, channel state identification followed by the appropriate utilization of NLOS range estimation is more beneficial.

In summary, these measurement results indicate that from the perspective of localization accuracy (i) NLOS identification can prove beneficial, as treating the NLOS range estimates differently from LOS range estimates can result in considerable reduction in the localization error, and (ii) the delay-spread statistics can be used successfully for NLOS identification in indoor UWB localization applications.

6.3.3 Propagation of Localization Error

In this section, we quantify using measurements, the extent of the propagation of error when sequential location estimation is applied, and demonstrate the efficacy of the error mitigation

algorithm discussed in Section 5.6. In order to utilize the available measurements to quantify the propagation of error with sequential location estimation, the locations of the transmit antenna are no longer considered as points in the "mobile trajectory", but as the locations of unlocalized nodes. Therefore, in this case, we have $N = 71$ unlocalized nodes within the area of interest, and their locations are sequentially estimated, starting with the node closest to the origin.

In our measurement setup, the transmitter was placed at one of $N = 71$ measurement locations, and the receiver at various anchor locations. Evidently, the main issue with using the available measurements for sequential location estimation, is the lack of range estimates between unlocalized nodes. In order to quantify the extent of the propagation of error in sequential location estimation, range estimates between transmit locations are required. Since these measurements are not available, we generate range estimates based on the statistical models discussed in chapter 2, using the parameters listed in Table 2.1. Specifically, LOS range estimates are modeled as unbiased Gaussian estimates of the true distances, and NLOS range estimates as unbiased Gaussian estimates of the true distance, superimposed with an exponentially distributed bias. Further, the probability that a given link between unlocalized nodes is NLOS is denoted by p_{NLOS}, and the transmission range is denoted by $R_{\max}$.

As discussed in Chapter 5, with regard to the availability of anchor nodes, there are two possibilities: (a) an infrastructure-based network, where range estimates from a large number of anchors deployed within the area of interest are available, and (b) an ad hoc network, where range estimates from a few anchors are located close to the origin are available, and location-awareness is sequentially propagated through the network. In order to evaluate the propagation of error in these two scenarios, we consider the following cases:

Case I: Sequential Localization in Infrastructure-based networks In this case, range estimates from $N_A = 18$ anchor nodes are assumed to be available at various locations, and the locations of nodes are sequentially estimated (from index $i = 1$ through $i = 71$) using (a) the LS-(Pure LOS) estimator, and (b) the LP-basic estimator. Figure 6.8 shows the estimated node locations and the corresponding root localization error of the two sequential estimators as the probability p_{NLOS} is increased. We see that due to the presence of anchors within the area of interest, we do not observe considerable propagation of error. This observation validates the analysis of the impact of geometry of available anchors on the propagation of error discussed in Section 5.5. The LP-basic sequential estimator again outperforms the LS-(Pure LOS) estimator in terms of average localization accuracy. The relative gain of using the LP-based approach over the LS-based approach in terms of average root localization error is approximately 33%. Further, we see that as the probability p_{NLOS} increases, the localization accuracy degrades. However, we note the performance of the LS-(Pure LOS) estimator is considerably better than in Figure 6.3(a), due to the availability of a larger number of LOS range estimates.

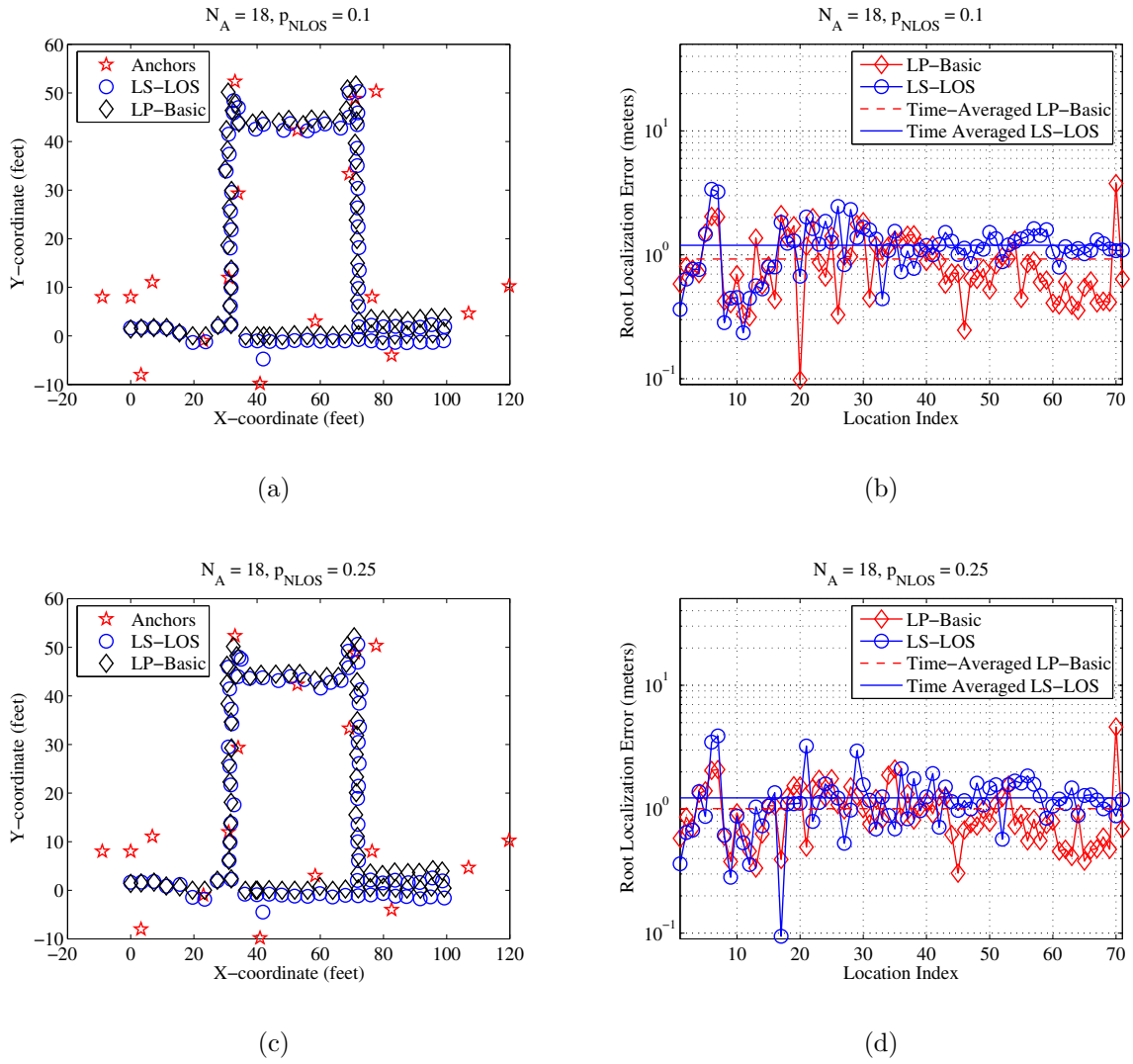


Figure 6.8: The performance of sequential estimation when $N_A = 18$ ("Infrastructure-based") using the LS-(Pure LOS) and LP-basic estimators when (a), (b) $p_{NLOS} = 0.1$, and (c), (d) $p_{NLOS} = 0.25$. The value of $R_{\max}$ used here is 10 meters.

Case II: Sequential Localization in Ad hoc networks In this case, $N_A = 5$ anchors are located around the origin, and the locations of the unlocalized nodes are estimated sequentially. Figures 6.9(a) and 6.9(b) illustrate that the localization error in such a scenario can propagate considerably. We see that the propagation of error can result in large localization error. However, we note that due to the accuracy of LOS range estimates, the overall structure of unlocalized nodes is still maintained, but the orientation of this structure is very sensitive to changes in the localization error as the distance from the anchors increases.

Figures 6.9(c) and 6.9(d) show the performance of the LP-based propagation mitigation ("'LP-PM"') method discussed in Section 5.6 that uses both NLOS range estimates and 2-hop range estimates to bound the location of the unlocalized node. We see that the application of the LP-PM method considerably limits the propagation of error, and displays significant gains when compared to the LS-(Pure LOS) method. For instance, when $p_{NLOS} = 0.1$, the average root localization error decreases from 8.7 meters when the LP-basic estimator is used to 2.4 meters when the LP-PM method is utilized. The root localization error achieved by the LP-PM method is larger than 1 meter, and further constraints can be added to improve localization accuracy at the cost of computational complexity. In general, as the local map of node locations remains intact, creating a feasible region for node locations can considerably limit the propagation of error by "guiding" the local map to the true locations. Physical constraints of the area of interest can be used to further reduce the size of the feasible region. This example emphasizes the utility of optimization methods in ad hoc location estimation.

6.4 Conclusions

In this chapter, we presented results from indoor UWB measurements which demonstrated the efficacy of the proposed NLOS identification, NLOS mitigation, and propagation of error mitigation algorithms. The application of the proposed NLOS identification algorithms based on received signal statistics to the available measurements indicated a low probability of identification error and negligible degradation in localization accuracy relative to perfect channel information. The LP-based NLOS mitigation algorithms proposed in Chapter 4 successfully incorporate NLOS range estimates into location estimation, outperform LS-based location estimators, and allow us to estimate node locations to within a root localization error of 1 meter. The propagation of localization error was found to be inconsequential when sequential location estimation is applied to infrastructure-based PoLoNets due to the presence of anchors within the area of interest resulting in superior anchor geometry. On the other hand, we verified that in ad hoc PoLoNets, there is considerable propagation of localization error when sequential estimation approaches are applied. The proposed LP-based confinement of a node's location using multi-hop and NLOS range estimates was found to considerably limit the propagation of localization error. The inclusion of additional constraints and optimization methods can further improve localization accuracy in PoLoNets.

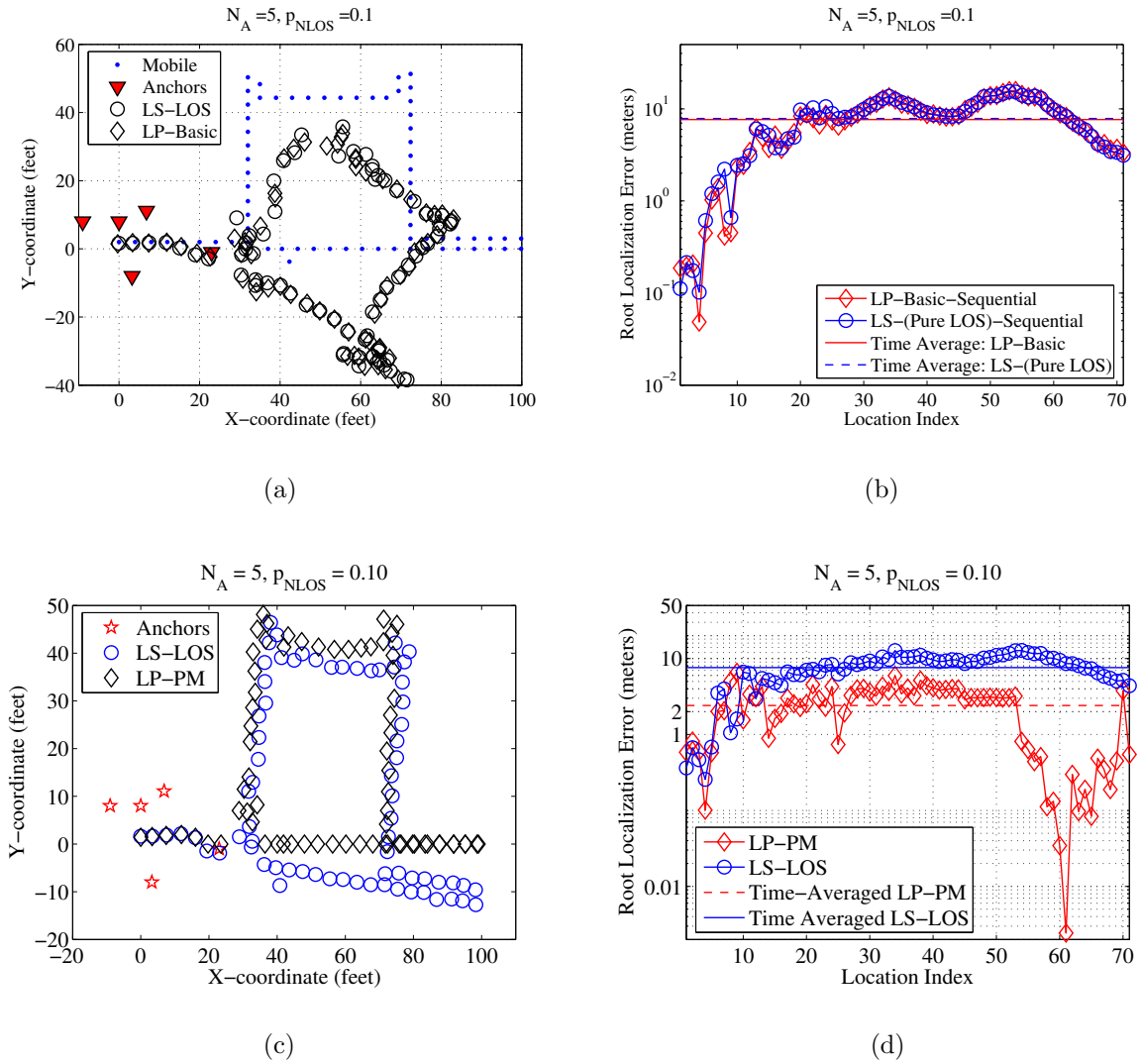


Figure 6.9: The performance of sequential estimation when $N_A = 5$ ("ad hoc"), $p_{NLOS} = 0.1$, (a), (b) using the LS-(Pure LOS) and LP-basic estimators, and (c), (d) when the LP-based error propagation mitigation method ("LP-PM") discussed in Section 5.6 was applied.

6.5 Appendix 6A: Kalman Smoothing

Suppose we are given noisy location estimates $\hat{\boldsymbol{x}}(t) = [\hat{x}(t)\ \hat{y}(t)]^T$, at discrete N instants of time $t = t_0 + k\Delta_t$, $k = 1, 2, \cdots, N$, where t_0 is the initial measurement time and Δ_t is the interval between the successive attempts at estimation a node's location. Then these estimates can be smoothed using a Kalman filter, so as the limit the sudden increase in localization error at any instant k based on previous location estimates.

The state vector is defined as

$$\boldsymbol{X}(t_k) = [x(t_k)\ y(t_k)\ v_x(t_k)\ v_y(t_k)]^T, \tag{6.1}$$

where $x(t_k)$ and $y(t_k)$ respectively represent the node's x and y coordinates, and $v_x(t_k)$ and $v_y(t_k)$ denote the velocities in the x and y directions at time t_k. Under the Kalman filtering framework, the state vector is assumed to follow the linear recursion:

$$\boldsymbol{X}(t_{k+1}) = \boldsymbol{\Phi}\boldsymbol{X}(t_k) + \boldsymbol{\Gamma}\boldsymbol{w}(t_k),\ k = 1, 2, \cdots, N, \tag{6.2}$$

where

$$\boldsymbol{\Phi} = \begin{bmatrix} 1 & 0 & \Delta_t & 0 \\ 0 & 1 & 0 & \Delta_t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},\ \boldsymbol{\Gamma} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta_t & 0 \\ 0 & \Delta_t \end{bmatrix}. \tag{6.3}$$

The entries of $\boldsymbol{w}(t_k)$ are assumed to independent Gaussian random variables:

$$\boldsymbol{w}(t_k) = [w_x(t_k) w_y(t_k)]^T \sim \mathcal{N}(0, \boldsymbol{Q}), \tag{6.4}$$

with $\boldsymbol{Q} = \sigma_w^2 \boldsymbol{I}_2$.

The interpretation of the above model is as follows. If the node is located at $\boldsymbol{x}(t_k) = [x(t_k)\ y(t_k)]^T$ with velocity vector $\boldsymbol{v}(t_k) = [v_x(t_k)\ v_y(t_k)]^T$ at time t_k, then after time Δ_t, it is assumed to have moved to position $\boldsymbol{x}(t_{k+1}) = [x(t_k) + \Delta_t v_x(t_k)\ y(t_k) + \Delta_t v_y(t_k)]^T$. The components of the actual velocity vector are now changed by a random amount $\Delta_t \boldsymbol{w}(t_k)$. Thus the total acceleration at time t_k is a modeled as a Rayleigh random variable:

$$a(t_k) = \sqrt{w_x^2(t_k) + w_y^2(t_k)},\ E\left\{a(t_k)\right\} = \sigma_w \sqrt{\frac{\pi}{2}}. \tag{6.5}$$

Since $\boldsymbol{X}(t_k)$, $k = 1, 2, \cdots, N$ contains the information of interest, only noisy observations of the location estimates are available, i.e.,

$$\boldsymbol{Y}(t_k) = \boldsymbol{M}\boldsymbol{X}(t_k) + \boldsymbol{U}_k, \tag{6.6}$$

where

$$\boldsymbol{M} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\ \boldsymbol{U}_k \sim \mathcal{N}(0, \boldsymbol{R}). \tag{6.7}$$

The matrix $\boldsymbol{R}$ needs to be estimated from measurements with known locations in a test scenario.

From the Gaussian observations, the minimum variance estimate of state vector is given by the conditional expectation:

$$\hat{\boldsymbol{X}}(t_k) = E\left\{\boldsymbol{X}(t_k)|\tilde{\boldsymbol{Y}}(t_k)\right\}, \tag{6.8}$$

where $\tilde{\boldsymbol{Y}}(t_k) = \{\boldsymbol{Y}(t_0), \boldsymbol{Y}(t_1), \cdots, \boldsymbol{Y}(t_{k-1}), \boldsymbol{Y}(t_k)\}$ represents the set of all current and previous observations. The optimal predicted value of the state vector $\boldsymbol{X}(t_k)$ given $\tilde{\boldsymbol{Y}}(t_{k-1})$ is given by:

$$\begin{aligned}
\hat{\boldsymbol{X}}_{k-1}(t_k) &= E\left\{\boldsymbol{X}(t_k)|\tilde{\boldsymbol{Y}}(t_{k-1})\right\} \\
&= E\left\{\boldsymbol{\Phi}\boldsymbol{X}(t_{k-1}) + \boldsymbol{\Gamma}\boldsymbol{w}(t_{k-1})|\tilde{\boldsymbol{Y}}(t_{k-1})\right\} \\
&= \boldsymbol{\Phi}E\left\{\boldsymbol{X}(t_{k-1})|\tilde{\boldsymbol{Y}}(t_{k-1})\right\} = \boldsymbol{\Phi}\hat{\boldsymbol{X}}_{k-1}(t_{k-1}).
\end{aligned} \tag{6.9}$$

The corresponding covariance matrices are given by

$$\boldsymbol{C}_k(t_k) = Cov\left\{\boldsymbol{X}(t_k)|\tilde{\boldsymbol{Y}}(t_k)\right\}, \tag{6.10}$$

and

$$\begin{aligned}
\boldsymbol{C}_{k-1}(t_k) &= Cov\left\{\boldsymbol{X}(t_k)|\tilde{\boldsymbol{Y}}(t_{k-1})\right\} \\
&= E\left\{\left(\hat{\boldsymbol{X}}_{k-1}(t_k) - \boldsymbol{X}(t_k)\right)\left(\hat{\boldsymbol{X}}_{k-1}(t_k) - \boldsymbol{X}(t_k)\right)^T \middle|\tilde{\boldsymbol{Y}}(t_{k-1})\right\} \\
&= \boldsymbol{\Phi}\boldsymbol{C}_{k-1}(t_{k-1})\boldsymbol{\Phi}^T + \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^T.
\end{aligned} \tag{6.11}$$

Optimal recursive estimators of minimum variance are obtained by the Kalman-Bucy filter defined in the following theorem:

Theorem 8. *The minimum variance estimator of the state of the system at time t_k is given by*

$$\hat{\boldsymbol{X}}_k(t_k) = \hat{\boldsymbol{X}}_{k-1}(t_k) + \boldsymbol{K}(t_k)\left(\boldsymbol{Y}(t_k) - \boldsymbol{M}\hat{\boldsymbol{X}}_{k-1}(t_k)\right). \tag{6.12}$$

Covariance matrices are updated using

$$\boldsymbol{C}_k(t_k) = \boldsymbol{C}_{k-1}(t_k) - \boldsymbol{K}(t_k)\boldsymbol{M}\boldsymbol{C}_{k-1}(t_k), \tag{6.13}$$

where

$$\boldsymbol{K}(t_k) = \boldsymbol{C}_{k-1}(t_k)\boldsymbol{M}^T\left(\boldsymbol{M}\boldsymbol{C}_{k-1}(t_k)\boldsymbol{M}^T + \boldsymbol{R}\right)^{-1} \tag{6.14}$$

is the Kalman gain. With initial values $\hat{\boldsymbol{X}}_0(t_0)$ and $\boldsymbol{C}_0(t_0)$ the recursion (6.12) can be evaluated via (6.9), (6.11), (6.13), and (6.14). The corresponding algorithm is described

below, where updated values are denoted by $^+$. *Let* $\boldsymbol{X}^+ = \hat{\boldsymbol{X}}_0(t_0)$ *and* $\boldsymbol{C}^+ = \boldsymbol{C}_0(t_0)$. *Then iterate:*

$$\begin{aligned}
\boldsymbol{X} &= \boldsymbol{X}^+, \\
\boldsymbol{C} &= \boldsymbol{C}^+, \\
\overline{\boldsymbol{C}} &= \boldsymbol{\Phi}\boldsymbol{C}\boldsymbol{\Phi}^T + \boldsymbol{\Gamma}\boldsymbol{Q}\boldsymbol{\Gamma}^T, \\
\boldsymbol{K} &= \overline{\boldsymbol{C}}\boldsymbol{M}^T\left(\boldsymbol{M}\overline{\boldsymbol{C}}\boldsymbol{M}^T + \boldsymbol{R}\right)^{-1}, \\
\boldsymbol{C}^+ &= \overline{\boldsymbol{C}} - \boldsymbol{K}\boldsymbol{M}\overline{\boldsymbol{C}}, \\
\boldsymbol{X}^+ &= \boldsymbol{\Phi}\boldsymbol{X} + \boldsymbol{K}\left(\hat{\boldsymbol{x}}_k - \boldsymbol{M}\boldsymbol{\Phi}\boldsymbol{X}\right).
\end{aligned}$$

A reasonable initial choice of $\hat{\boldsymbol{X}}_0(t_0)$ is $[\hat{\boldsymbol{x}}(t_0)^T 0\ 0]^T$, i.e., the speed of the mobile is assumed to be zero initially. Based on the above theorem, the state vector at any time instant t_k can be predicted, based on the observations (computed location estimates) in the previous $k-1$ time instants.

Chapter 7

MAC design for UWB PoLoNets

7.1 Introduction

In communication networks, the physical channel is often shared by several communicating devices or users. A packet radio network is an example of such a multiple-access system where different users (or nodes) attempt to communicate using the common wireless channel. A transmission by any node is likely to be heard by several other nodes and therefore can distort intended transmissions to those nodes. Therefore, we require a mechanism or protocol through which the channel is divided or shared among the users. Such a protocol is called a Medium-Access Control (MAC) protocol and is implemented in the MAC sublayer of the Data Link Control (DLC) layer of the ISO/OSI network model [80]. The function of the MAC sublayer is to allocate the multiple-access channel to the users such that each user can transmit its frames (packets) without undue interference from other users.

In this chapter, we first discuss the required characteristics of a MAC protocol for UWB PoLoNets. We then use results derived in Chapter 3 to establish a connection between localization accuracy and the selection of MAC protocols. Based on this connection, we propose a spread-spectrum multiple-access scheme for UWB PoLoNets which is shown to outperform the traditional Carrier-Sense Multiple-Access (CSMA) protocols in terms of the average localization accuracy achieved.

Original Contributions in Chapter 7:

- The specification of the desired features of MAC protocols in UWB PoLoNets,
- Establishing the connection between MAC protocol design and localization accuracy through the properties of bounds on location estimation.
- The specification of a modified spread-spectrum MAC protocol [23] for distributed ranging in ad hoc UWB PoLoNets, by comparison the traditional CSMA protocol via simulations.

The organization of this chapter is as follows: In Section 7.2, we discuss the desired features of MAC protocols in ad hoc UWB PoLoNets. Section 7.3 provides a detailed description of a spread-spectrum MAC protocol developed specifically for ad hoc UWB PoLoNets. Simulation results comparing the performance of the proposed multiple-access scheme with a CSMA-based MAC protocol in terms of localization accuracy are presented in Section 7.4. Our conclusions are presented in Section 7.5.

7.2 MAC Design for ad hoc UWB PoLoNets

In Chapter 1, we made the observation that a PoLoNet can be viewed as the hybrid of a mobile ad hoc network and a sensor network. Specifically, the data sinks of the network are the mobile and command-and-control nodes, and the physical quantity being measured through the stationary reference nodes is the physical location of the mobile nodes. In the case of ad hoc PoLoNets, the size of the network is not fixed *a priori*, and consequently, network scalability is essential. In addition, the MAC protocol designed for PoLoNets needs to provide automated node discovery, fair and reliable multi-hop data transfer and routing between nodes, and possess the ability to adapt to mobility and node failure. Further, the requirement of up-to-date knowledge of mobile locations implies constraints on latency need to be satisfied. Evidently, there is a trade-off between update-rate and reliability of location estimates, and a useful feature of the MAC protocol would be the flexibility to trade one quantity for the other. When the desired PoLoNet lifetimes are large, the energy-efficiency of the MAC protocol may be an important consideration. Another constraint that needs to be considered is the hardware complexity of the nodes. In order to minimize the hardware complexity of the nodes, we may need to enforce the constraint that at any instant, a node can either transmit or receive, but not both. Additionally, each node is required to "listen" on only one channel at any time.

7.2.1 MAC Design and localization error

In Section 3.3, we showed that on the average, the accuracy of the LS location estimate of an unlocalized node is determined by the number of ranges successfully received from other localized nodes. It was shown using the CRLB that, except when all localized nodes were collinear, increasing the number of range estimates results in reduction of the localization error. Even when connectivity with localized nodes is limited, repeated range measurements allow averaging of range estimates, which reduces their variance and hence the localization error.

Therefore, a MAC protocol that allows each unlocalized node to accumulate several range estimates in a short duration increases the likelihood that an accurate estimate of the nodes location is computed at the end of that duration. In the case of reference nodes, these

accurate location estimates subsequently translate to accurate estimates of the location of the mobile nodes, since mobile nodes utilize these coordinates to triangulate their own locations. In the case of mobile nodes, the lower latency in obtaining accurate location estimates allows mobile nodes to obtain estimates of their locations more frequently, which is especially important at high mobile node speeds.

In terms of the ranging scheme discussed, this implies that a MAC protocol which provides a higher effective throughput of range estimates (RI and RR packets) allows (i) faster formulation of an initial location estimate, and (ii) faster convergence of location estimates of unlocalized nodes to their true locations. This conjecture is analytically examined in Chapter 8. Similarly, in the second phase of the PoLoNet evolution, where the estimates of locations of mobile nodes are routed back to a monitoring station, a higher effective throughput reduces the lag between the estimation of the mobiles location, and the availability of this information at a remote monitoring location (provided the MAC protocol treats range-packet handshakes and data exchanges in the same manner).

7.2.2 UWB Physical Layer

Due to stringent constraints on the transmit power, data-rates are reduced in order for UWB radios to achieve significant transmission ranges in indoor environments. These low-data rates over a wide transmission bandwidth implies large spreading gains (bandwidth to data-rate ratio) which indicates that "single-channel" approaches are wasteful in terms of throughput. In order to take advantage of the inherent spreading in UWB channels, and to lower the transmit power spectral densities, spread-spectrum techniques such as Time-Hopped (TH) or Direct Sequence (DS) spreading are expected to be an integral feature [13], [35] of the signals used.

While the majority of the MAC protocols proposed for mobile ad hoc and sensor networks are based on CSMA [81], there are several issues with CSMA-based schemes for networks with a UWB physical layer. With the use of spread-spectrum signaling, in order for a node sensing the channel to detect channel usage, the spreading code used for transmission must be known beforehand. As a result, detection of the UWB signal after spreading is likely to be highly computationally intensive and time-consuming, i.e., simple energy detection is not sufficient. This implies that reliable sensing of the channel can be achieved only through the acquisition of the transmitted signal. While CSMA-based MAC protocols are scalable, pure CSMA suffers from the Hidden-Node problem, and we need RTS/CTS overheads to alleviate this problem, and this additional overhead reduces the energy efficiency of the network. TDMA is a commonly suggested MAC protocol [82] in sensor network literature as it is power efficient and allows for sleep schedules to improve power efficiency, but is not scalable and requires network wide synchronization.

A multi-channel spread-spectrum protocol allows simultaneous transmissions by the nodes of the network at the cost of incurring multi-access interference; a single-channel approach such

as CSMA on the other hand does not allow nodes in the same vicinity to transmit simultaneously. Based on the above reasoning, a spread-spectrum MAC protocol could provide a higher localization accuracy in UWB PoLoNets than CSMA-based protocols. The details of a spread-spectrum multiple-access protocol, designed specifically for UWB PoLoNets are provided in the following section.

7.3 Proposed Solution: Spread-Spectrum MAC Protocol

As discussed in 2.3, in the absence of network-wide synchronization, the process of ranging between two nodes is performed through an exchange [18] of RI and RR packets. An efficient MAC protocol from the perspective of localization accuracy (i) provides higher rates of packet exchanges within the network between unlocalized and localized nodes, and (ii) is scalable, allowing packet exchanges between an arbitrary number of unlocalized nodes and localized nodes. The MAC protocol we propose for ad hoc UWB position-location networks is based on the Common-Transmitter Spread Spectrum Multiple-Access protocol (CT-SSMA) [30], modified to in order meet the above requirements.

The details of the scheme are as follows: when node location-estimates are either unknown or not updated (in the case of mobile nodes), nodes would ideally like to obtain range-estimates from every other localized node in the network. The contention between such unlocalized nodes for broadcasting RI packets to localized nodes occurs on a "common" [30] spreading code C_0 and all nodes, when idle, are listening on the common code C_0. However, each RI packet specifies a "private" code C_i, unique to the node broadcasting it, on which RR packets from localized nodes are to be received. As discussed below, this allows several range exchanges to proceed simultaneously in all regions of the PoLoNet, at the cost of increased multiple-access interference.

7.3.1 Multiple-Access Ranging

The RI packet broadcast by a node i contains[1] a new TH-code C_i. As soon as node i transmits the RI packet on code C_0, it begins to listen on code C_i for a window of time from $t = 0$ to $t = T_W$. The window-length T_W is much larger than the duration T_{RR} of an RR packet, $T_W \gg T_{RR}$, which allows multiple RR packets to be gathered within the duration of

[1] In order to avoid large overheads, the RI packet from node i could specify a parameter that can be used to uniquely compute C_i. For instance, in the case of random codes, the seed used to generate the random TH (or DS) sequences can be specified. Further, the use of random codes does not place a stringent limitation on the number of available private codes.

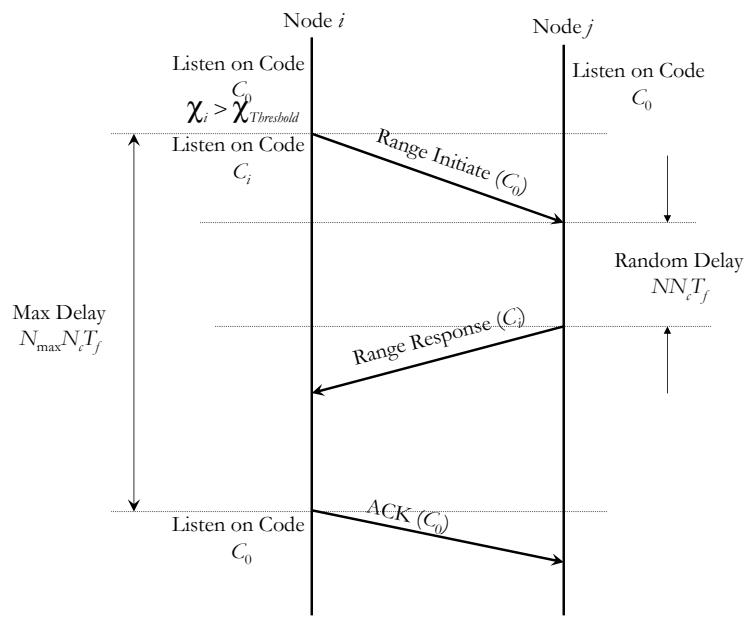


Figure 7.1: Timing Diagram for the Proposed MAC protocol.

the window[2].

In order to allow an arbitrary number of range estimates to be gathered within this window, the following scheme, illustrated in Figure 7.1 is used: suppose node i broadcasts the RI packet on C_0 at time $t = t_1$. In the absence of collisions, this packet reaches a localized node j at $t = t_1 + \tau_0$, where $\tau_0 = \frac{R_{ij}}{c}$ represents the propagation delay between nodes i and j. Node j synchronizes itself to the TOA of the received packet. Node j then responds with an RR packet on the private code C_i after a random integer multiple of a known delay Δ_T (at $t = t_1 + \tau_0 + k\Delta_T$), where k is a random positive integer in the interval $\left[1, \left\lfloor \frac{T_W}{\Delta_T} \right\rfloor\right]$. Node i receives the RR packet, again in the absence of collisions, and determines the TOA of the RR packet from node j, which is $t_2 = 2\tau_0 + t_1 + k\Delta_T$. Node A then computes the difference $\Delta t = t_2 - t_1 = 2\tau_0 + k\Delta_T$, from which the known delay Δ_T can be eliminated, without any ambiguity in interpreting the range estimate, by ensuring that $\tau_0 \ll \Delta_T$. The range estimate r_{ij} can then be computed using $r_{ij} = \frac{c(t_2 - t_1 - k\Delta_T)}{2}$ after processing delays have be calibrated out.

The use of the random delay $k\Delta_T$, $k \in \left[1, \left\lfloor \frac{T_W}{\Delta_T} \right\rfloor\right]$, in the above mechanism is for collision-avoidance between RR packets transmitted on the private code of the unlocalized node that

[2]The duration of the window T_W can be adapted based on the desired localization-accuracy and acceptable latency, since a longer window-length will likely result in a larger number of received RR packets. However, this discussion is outside the scope of this work.

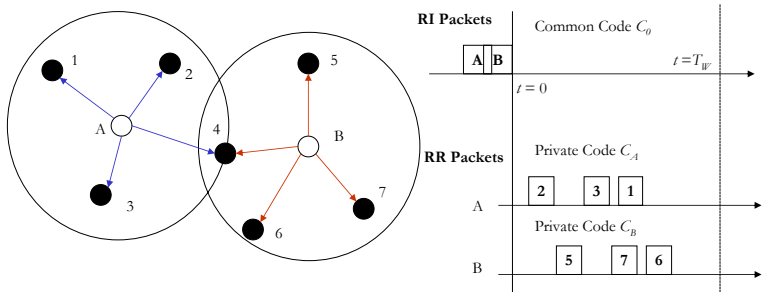


Figure 7.2: An illustration of the multiple-access ranging procedure. In this situation, two nodes A and B simultaneously wish to obtain range estimates from localized nodes around them (nodes 1 through 7). The RI packets broadcast by nodes A and B on the common code C_0 collide at node 4, and as a result node 4 does not respond with RR packets. The remaining nodes respond with RR packets on the private codes C_A and C_B, each offset by a random delay $k\Delta_T$. Nodes A and B each collect $m = 3$ range estimates at the end of this process.

transmitted the RI packet. Once node j transmits its RR packet, it continues listening on code C_0. The unlocalized node i decodes data on code C_i and collects the received packets from all localized nodes that send an RR packet. An RR packet can be assumed to be successfully received if the SINR at node i during the time of reception exceeds the threshold $\xi_{\min}$. Node i thereby estimates the distance r_j to every localized node j that sends an RR packet, and computes an estimate $\hat{\boldsymbol{x}}_i$ of its location. The accuracy of the range estimate r_j is also determined by the SINR at node i through (2.25). At $t = T_W$, node i assumes that all other localized nodes are out of range and broadcasts an ACK packet on C_0 declaring the newly acquired range estimates and the estimated coordinates. An example scenario based on the above protocol mechanism is illustrated in Figure 7.2, where two unlocalized nodes simultaneously attempt to receive range estimates.

If node i receives no RR packets (potentially due to (a) collision between RI packets on the common code C_0, (b) collisions between RR packets on the private code C_i, or (c) because there are no localized nodes within range) within a certain interval, it times out and (exponentially) backs-off for $T_W \cdot 2^{n-1}$ seconds, where n represents the index of the retransmission attempt. Additionally, node i can increase the transmit power from P_i to $P_i + \Delta P$ in its subsequent retransmission attempt. If node i receives more RR packets within the window than necessary to maintain the localization accuracy, it then decreases the transmit power from P_i to $P_i - \Delta P$ in its subsequent range-initiate transmission, which reduces power consumption. Such power-control algorithms based on localization accuracy [22] are the subject of Chapter 9.

Due to the ability of the discussed ranging mechanism to "piggyback" data on RI and RR packets, a node that transmits a RI packet can transmit all the range and coordinate

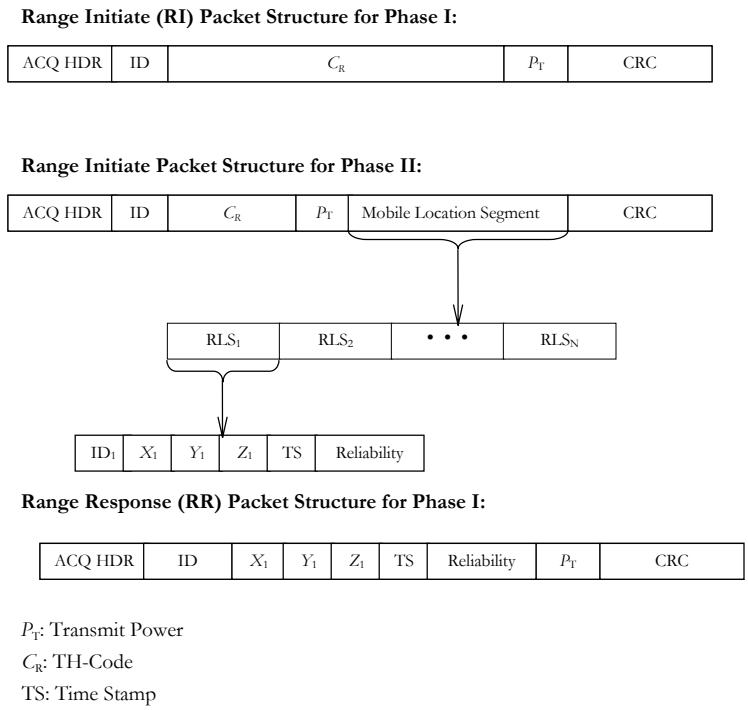


Figure 7.3: Range-Initiate (RI) and Range-Response (RR) Packet structures for the two phases of the network. Note that the transmit power levels are included in each RI packet; if the node that receives this packet wants to communicate/range with the node that transmitted this RI packet, it knows what transmit power-level to use. This is useful when we implement power-control algorithms for the network.

information it has along with the time-stamps for each of these values. It must be noted that the same MAC protocol can be used for data transmissions in the second phase of the network, to route the coordinates of the nodes to a data sink, by making a distinction between the data and ranging packets. In such a case, the relatively high throughput promised by the proposed approach result in a lower latency in the arrival of this data at the data-sink that monitors the node locations.

A possible structure for the transmitted packets in the two network phases is shown in Figure 7.3. In this way, each node that listens to this packet obtains all these values and possibly can use these values to obtain/improve location estimates. The benefits of the proposed multiple-access mechanism are as follows:

- The protocol takes advantage of the inherent spreading present in the UWB physical layer;
- It allows independent ranging and communication to proceed independently and si-

multaneously in different parts of the network (spatial reuse) and therefore reduces latency;

- The protocol is scalable, with a larger number of nodes being accommodated at the price of higher multiple-access interference, leading to graceful degradation;
- Network-wide synchronization is not required;
- The hardware complexity of a node is kept small by ensuring that it is not necessary for a node to transmit and receive packets at the same time, or receive packets on two spreading codes simultaneously;
- Power control can be used to achieve reliability versus energy efficiency trade-offs by controlling the number of neighboring nodes that can hear a range-initiate packet.

A potential disadvantage of the proposed scheme results in the presence of clock drift, as discussed in Section 2.3. If the duration of the window T_w is large, this can lead to considerable range estimation error determined by (2.16). In the following section, we present simulation results comparing the performance of the proposed multiple-access scheme with CSMA.

7.4 Simulation Results

In order to evaluate the performance of the proposed multiple-access scheme, we compare its performance with carrier-sense multiple-access (CSMA), as CSMA forms the basis of most MAC protocols typically investigated for mobile ad hoc and sensor networks. The performance of the proposed scheme was compared with CSMA in terms of the average convergence rate of the total localization error of mobile and reference node location estimates, as well as the average number of RR packets successfully received in an interval. The total localization error of the PoLoNet at any instant t is defined as

$$\Omega(t) = \frac{\left(\overbrace{\sum_{i=1}^{N_M} \|\boldsymbol{x}_i(t) - \hat{\boldsymbol{x}}_i(t)\|^2}^{mobile\ nodes} + \overbrace{\sum_{k=1}^{N_R} \|\boldsymbol{x}_k - \hat{\boldsymbol{x}}_k(t)\|^2}^{stationary\ reference\ nodes} \right)}{N_M + N_R}.$$

In the CSMA case, RI and RR packets are transmitted on a single TH-code C_0. Nodes "sense" the channel by listening on C_0 before transmitting. A transmission by a node j is sensed at node i if the SNR ξ corresponding to that transmission exceeds a threshold ξ_L at node i. A node that transmits an RI packet awaits RR packets from localized nodes

within a window of duration T_W. If no RR packets are received within that window (or if an ongoing transmission was sensed before broadcasting the RI packet), the node exponentially backs-off for $T_W \cdot 2^{n-1}$ seconds, where n represents the index of the retransmission attempt, before re-broadcasting the RI packet.

In these simulations, the reference and mobile nodes are deployed at $t = 0$ and the mobile nodes move linearly with speed $v_M = 0.25\ m/s$. It was assumed that $N_M = 2$, $N_R = 10$ and $N_A = 5$. The stationary reference nodes were uniformly distributed over the $L \times L$ area, defined by the lines $x = 0$, $x = L$, $y = 0$, $y = L$. The fixed anchor nodes were assumed to located outside the area of interest at $\{(L/2 \pm L/4, -L/2 \pm L/4), (L/2, -L/2)\}$.

The LS estimator defined by (3.16) was used to estimate node locations with the total localization error of unlocalized nodes arbitrarily set at a large value ($\sim 2000\ m^2$) at $t = 0$. The value of K_R used to generate the range estimate variances in (2.25) was $K_R = 100$, and the transmit power of all nodes was fixed at $P_t = 1\ mW$. The probability of an unlocalized node transmitting an RI packet in any given time-slot was set at $p = 0.15$, and the duration of the window $T_W = 20T_s$, where the width of a slot was $T_s = 50\ ms$, and the durations of RI and RR packets and the delay Δ_T are assumed to be equal: $T_{RI} = T_{RR} = \Delta_T = 20\ ms$.

The SINR at a node j corresponding to a signal from node i is computed using the relation

$$\xi_{j \leftarrow i} = \frac{K_P P_t R_{ij}^{-\beta}}{N_0 + \sum_{k \neq i} g_k K_P P_t R_{ik}^{-\beta}}, \tag{7.1}$$

where k represents the index of simultaneous transmissions, N_0 represents the power spectral density of additive noise, and g_k incorporates the effect of spreading gain if multiple TH codes are used. If the kth simultaneous transmission occurs on the same code as node i, then $g_k = 1$, else, $g_k = 1/64$. The SINR and SNR thresholds that determine the successful decoding of packets and the sensing of ongoing packet transmissions are $\xi_{\min} = 15$ dB and $\xi_L = 5$ dB respectively.

The performance of the two protocols in terms of the average number of RR packets received successfully within the network versus time for different path loss exponents was compared. Figure 7.4 depicts the performance for $\beta = 2$. We see that the proposed approach has a much higher effective throughput of RR packets. The average total localization error $E\{\Omega(t)\}$ was computed by averaging $\Omega(t)$ over a large number of simulation runs. Figures 7.5(a) and 7.5(b) compare the performance of the two protocols in terms of the average total localization error of mobile and reference nodes versus time for different values of β. We see that the average total localization error $E\{\Omega(t)\}$ decreases much faster with t for the proposed approach than for the CSMA scheme, which is expected since the localization error and the average effective RR packet throughput are strongly (negatively) correlated.

In the initial phase after deployment after $t = 0$, relatively few localized nodes (only N_A fixed anchors at $t = 0$) can provide range information to unlocalized (reference and mobile) nodes. Consequently, location estimates are inaccurate, and can even be outside the area

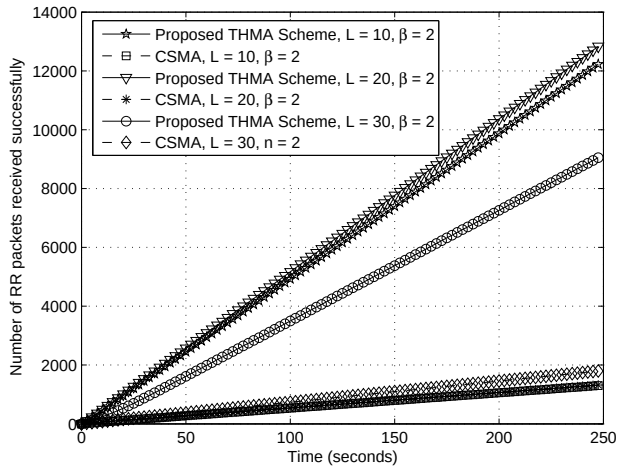


Figure 7.4: Comparison of CSMA and the proposed scheme in terms of the average number of successful RR packet transmissions versus time for $\beta = 2$. The values of K_P and K_R are 1×10^6 and 100 respectively.

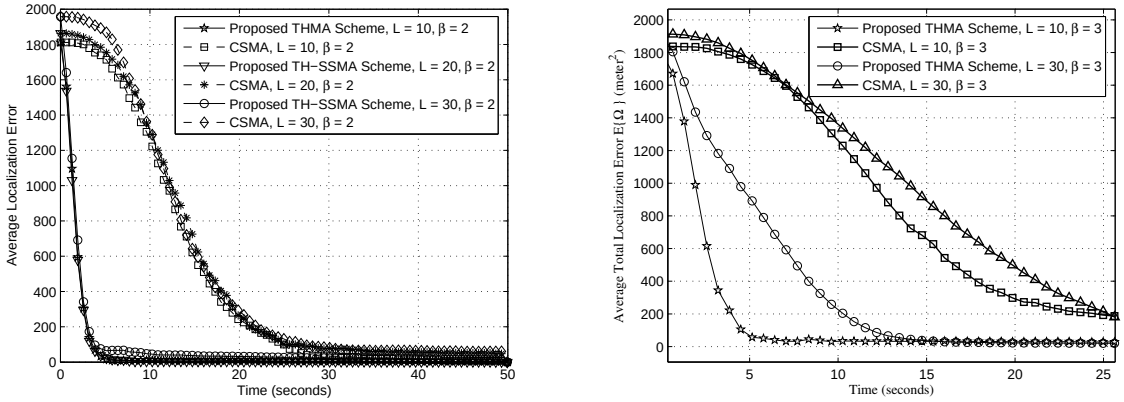
of interest[3] leading to large localization errors. As time progresses, more reference nodes estimate their locations and consequently, more range estimates are available to the mobile nodes and unlocalized reference nodes. At this point, the total average localization error decreases sharply.

With $\beta = 2$, an increase in the area (L^2) of the network does not have a significant impact on performance. For the $\beta = 3$ case[4], increasing the area of interest reduces the effective throughput of RR packets. There are two opposing factors at work: as the area increases, the average distance between nodes increases, increasing range estimate variances and possibly decreasing the probability of successful detection of RI and RR packets. However, as the area increases, the inherent spatial reuse increases; if the density of nodes is kept constant as the area of the network is increased, the benefits of spatial reuse in terms of increased effective throughput are observed.

Figure 7.6 shows the mobile locations and the corresponding estimates obtained through the proposed MAC protocol, in a $L \times L \times L$ volume of interest, $L = 20$ meters, with $N_A = 5$ anchors and $N_R = 20$ reference nodes. The components of the mobile speed in the x, y, and z directions are 0.2 meters per second. We see that in the initial segment of the mobile's trajectory, the localization accuracy of the mobile is poor as the number of localized nodes in the volume of interest in small. Once the reference nodes within the volume of interest

[3]It is assumed that the nodes are not aware of the boundaries of the area of interest, precluding the possibility of eliminating such egregious location-estimates.

[4]Although we assume a LOS model for the range estimation process, this scenario represents the case where the LOS multipath component is severely attenuated (soft-NLOS).



(a) (b)

Figure 7.5: Comparison of CSMA and proposed multiple-access scheme in terms of the total localization error (meter^2) versus time when (a) $\beta = 2$, and (b) $\beta = 3$. The values of K_P and K_R are 1×10^6 and 100 respectively.

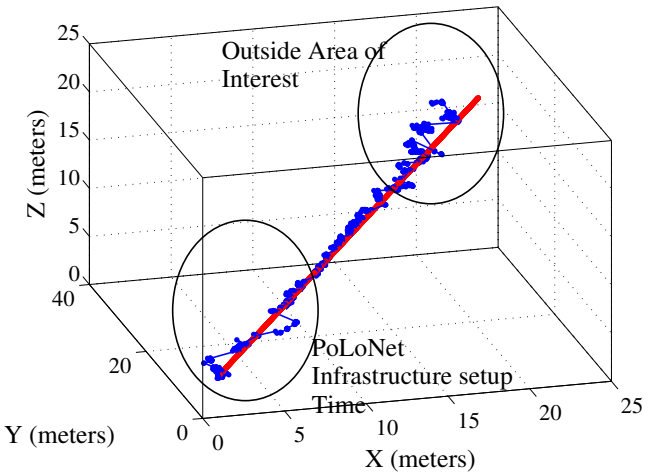


Figure 7.6: Mobile locations and estimates obtained via simulation of the proposed MAC protocol with a PoLoNet containing $N_R = 20$ reference nodes, and $N_A = 5$ anchors located outside the volume of interest. The mobile speed is $v = 0.2\sqrt{3}$ meters per second, and travels linearly across the volume of interest.

are localized, accurate estimates of the mobile's location are obtained. In the latter part of the mobile's trajectory, the mobile leaves the volume of interest, and this results in a poor geometry of localized reference nodes that provide range estimates to the mobile, resulting in low localization accuracy.

7.5 Conclusions

This chapter discussed the problem of multiple-access design for ad hoc UWB PoLoNets.The properties of location estimation discussed in Chapter 3 provided insight into the problem of MAC design, and indicate that the problem of minimizing localization error in PoLoNets is equivalent to the problem of maximizing effective throughput of range information within the network. Based on insights into the nature of the UWB physical layer, we developed a spread-spectrum protocol based on time-hopping for UWB PoLoNets, which is based on the common-transmitter spread-spectrum multiple-access approach. Network simulations were performed to compare the performance of the proposed protocol with CSMA, which is one of the most commonly investigated protocols for ad hoc and sensor networks. We showed that the proposed protocol is superior to CSMA in terms of the localization accuracy and convergence time of the location estimates. In the following chapter, we discuss the mathematical modeling of the relationship between network parameters and average localization accuracy.

Chapter 8

Multiple-Access Analysis

8.1 Introduction

In chapter 7, based on the properties of location estimation, we conjectured that a MAC protocol which provides a higher effective throughput of range estimates to unlocalized nodes should allow faster convergence of node location estimates to the true locations. In this chapter, we present an in-depth examination of the relationship between localization accuracy and the effective throughput of packets in a spread spectrum multiple-access PoLoNet. We develop an analytical framework for characterizing the dependence of localization accuracy on network parameters such as node densities, mobile speed and other multiple-access parameters in synchronous and asynchronous PoLoNets. The insights into the impact of these parameters on the localization accuracy provided by the developed analytical framework are then verified using simulation results. The main challenges in the analytical modeling of multiple-access PoLoNets are then discussed, along with possibilities for further research.

Previous Work included in this chapter: The distribution of the received power from a Poisson field of interferers has been derived in [31].

Original Contributions in this chapter:

- A framework for the analytical modeling of the relationship between multiple-access system parameters and the resulting average localization accuracy in synchronous and asynchronous PoLoNets.
- Insights into the impact of the density of localized nodes and speed of mobile nodes on average localization accuracy based on analysis, and verification via simulation results.

The organization of this chapter is as follows: in Section 8.2, we examine the relationship

between the effective throughput of packets at an unlocalized node and the resulting average localization accuracy as a function of time. Based on the network model discussed in Section 8.3, the expressions for the average localization accuracy of a node in terms of various system parameters in synchronous and asynchronous PoLoNets are derived in Sections 8.4 and 8.5 respectively. A discussion of the analytical insights gained from these expressions is presented, followed by the validation of these insights using simulation results in Section 8.6. Section 8.6 also lists some aspects of the modeling of PoLoNets that warrant investigation and require more sophisticated modeling strategies. Finally, Section 8.7 summarizes the main results contained in this chapter.

8.2 Average Localization Error versus Throughput

Consider a slotted packet transmission system where each node is allowed to transmit a single packet within a slot. Let us restrict our attention to a single unlocalized node successfully receiving packets from localized nodes with a constant throughput per slot η. Each successful reception of a packet is assumed to result in successful range estimation, based on the TOA of the packet within a slot. Consequently, as η represents the probability that an unlocalized node receives a range estimate in a given time-slot, starting at $t = 0$, the probability that m range estimates are accumulated by an unlocalized node at time $t = nT_s$, is given by

$$p_{n,m} = \binom{n}{m} \eta^m (1-\eta)^{n-m}.$$

Therefore, the average localization error in the nth time-slot (at time $t = nT_s$) $\overline{\Omega(n)}$ of the unlocalized node is given by

$$\overline{\Omega(n)} = E\left\{\Omega(nT_s)\right\} = \sum_{m=0}^{n} p_{n,m}\Omega_{\boldsymbol{x},m} = \sum_{m=0}^{\infty} \binom{n}{m} \eta^m (1-\eta)^{n-m} \Omega_{\boldsymbol{x},m}, \tag{8.1}$$

where $\Omega_{\boldsymbol{x},m}$ is the average localization error as a function of the number of range estimates. Ignoring the impact of geometry, we model the average localization error $\Omega_{\boldsymbol{x},m}$ based on (3.7):

$$\Omega_{\boldsymbol{x},m} = \begin{cases} \Omega_0, & m < m_0, \\ \frac{4\sigma^2}{m}, & m \geq m_0, \end{cases} \tag{8.2}$$

where Ω_0 is the uncertainty in the location of an unlocalized node, and m_0 is the minimum number of range estimates required to formulate a location estimate. The variance of range estimates σ^2 is assumed to be constant for simplicity[1]. Substituting for $\Omega_{\boldsymbol{x},m}$ in (8.1), we

[1]In reality, the variance of TOA-based range estimates is dependent on the distance between localized and unlocalized nodes. However, for mathematical tractability, we neglect this dependence and assume that the variance of range estimates from all localized nodes is a constant.

have

$$\overline{\Omega(n)} = \Omega_0 \sum_{m=0}^{m_0-1} \binom{n}{m} \eta^m (1-\eta)^{n-m} + 4\sigma^2 \sum_{m=m_0}^{n} \binom{n}{m} \eta^m (1-\eta)^{n-m} \left(\frac{1}{m}\right). \tag{8.3}$$

If σ^2 is *independent of the parameters that determine the average effective throughput per slot* η, then it is straightforward to show that the system parameters that maximize η also minimize the average localization error $\overline{\Omega(n)}$ in any time slot n. In order to justify this claim, assume that x is a given system parameter that impacts the throughput η. Then, we see that an extreme point of η with respect to x is also an extreme point of $\overline{\Omega(n)}$ with respect to x:

$$\frac{\partial \eta}{\partial x} = 0 \Rightarrow \frac{\partial \overline{\Omega(n)}}{\partial x} = \frac{\partial \overline{\Omega(n)}}{\partial \eta} \cdot \frac{\partial \eta}{\partial x} = 0.$$

Further, when $\Omega_0 \gg 4\sigma^2$ (this is typically true for UWB systems), we can show that (the proof is provided in Appendix 8A) the second derivative of $\overline{\Omega(n)}$ with respect to x is positive if the second derivative of η with respect to x is negative:

$$\frac{\partial^2 \eta}{\partial x^2} < 0 \Rightarrow \frac{\partial^2 \overline{\Omega(n)}}{\partial x^2} > 0.$$

A similar result holds when the variance of range estimates is not constant, but remain independent of η. This implies that when the range measurement variances are independent of the system parameters (and consequently the effective throughput), maximizing the effective throughput is tantamount to minimization of the localization error.

Figure 8.1 shows the simulated values of average localization error $\overline{\Omega(n)}$ computed through (8.3) for different values of the slot-index n and the effective throughput η. As discussed above, we see that for any n, increasing η results in the decrease of $\overline{\Omega(n)}$. Thus, when the range estimate variances are constant and independent of the throughput, the problem of minimizing the average localization error is identical to the commonly studied problem of maximizing the throughput of packets within the network. This is an intuitive result as the accuracy of location estimates was shown to be dependent on the *amount* of range information available in Theorem 5, as conjectured in Chapter 7.

To provide a concrete example, consider a synchronous PoLoNet with a slotted TDMA multiple-access scheme. Let us assume that in a given slot, a single anchor transmits a packet to a specific unlocalized node. The unlocalized node estimates the distance between the nodes based on the TOA of the packet within the slot. It can be reasoned that the range measurement variance in any slot is independent of the throughput of packets, as the variance of a range estimate in a given slot is fixed by the received SNR, and independent of the events occurring in other slots. To illustrate the case in point, consider a system parameter such as the total number of unlocalized nodes. Assume that the total number of unlocalized nodes is increased keeping the number of anchors fixed, and the widths of the packets and slots are maintained the same. It is evident that for a fixed slot duration,

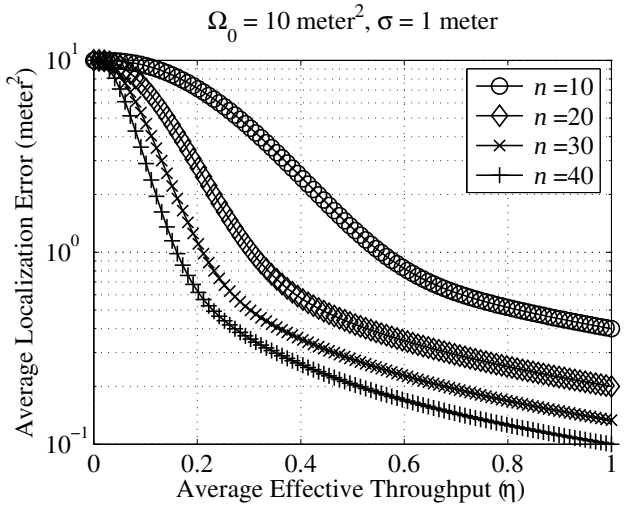


Figure 8.1: The variation of the average localization error $\overline{\Omega(n)}$ with average throughput η for different slot indices n. Here, $\Omega_0 = 10$ meter^2, and $\sigma = 1$ meter.

an increase in the number of unlocalized nodes reduces the average throughput of range estimates *per unlocalized node*. The range measurement variance would not be dependent on the number of unlocalized nodes, as multiple-access interference is absent. Thus, increasing the number of unlocalized nodes would reduce the *average* successful throughput per slot, thereby increasing the localization error as a function of time. It must be noted that an analysis along the above lines would be generally applicable for contention-free multiple-access schemes.

However, for contention-based schemes, where the channel is shared simultaneously by several nodes, the multiple-access interference simultaneously affects the effective throughput and the range estimation variance (through the SINR). Therefore, the choice of system parameters impacts both the throughput and the range measurement variances simultaneously as shown in Figure 8.2. Specifically, if x denotes a generic system parameter,

$$\begin{aligned}
\frac{\partial \overline{\Omega(n)}}{\partial x} &= \frac{\partial}{\partial x} \sum_{m=0}^{\infty} \binom{n}{m} \eta^m (1-\eta)^{n-m} \Omega_{\boldsymbol{x},m} \\
&= \sum_{m=0}^{\infty} \binom{n}{m} \left[\left(\frac{\partial}{\partial \eta} \eta^m (1-\eta)^{n-m} \right) \left(\frac{\partial \eta}{\partial x} \right) \Omega_{\boldsymbol{x},m} + \eta^m (1-\eta)^{n-m} \left(\frac{\partial \Omega_{\boldsymbol{x},m}}{\partial x} \right) \right] \\
&= \left(\frac{\partial \eta}{\partial x} \right) \sum_{m=0}^{\infty} \binom{n}{m} \left(\frac{\partial}{\partial \eta} \eta^m (1-\eta)^{n-m} \right) \Omega_{\boldsymbol{x},m} \\
&\qquad + \sum_{m=m_0}^{\infty} \binom{n}{m} \eta^m (1-\eta)^{n-m} \frac{4}{m} \left(\frac{\partial \sigma^2}{\partial x} \right).
\end{aligned}$$

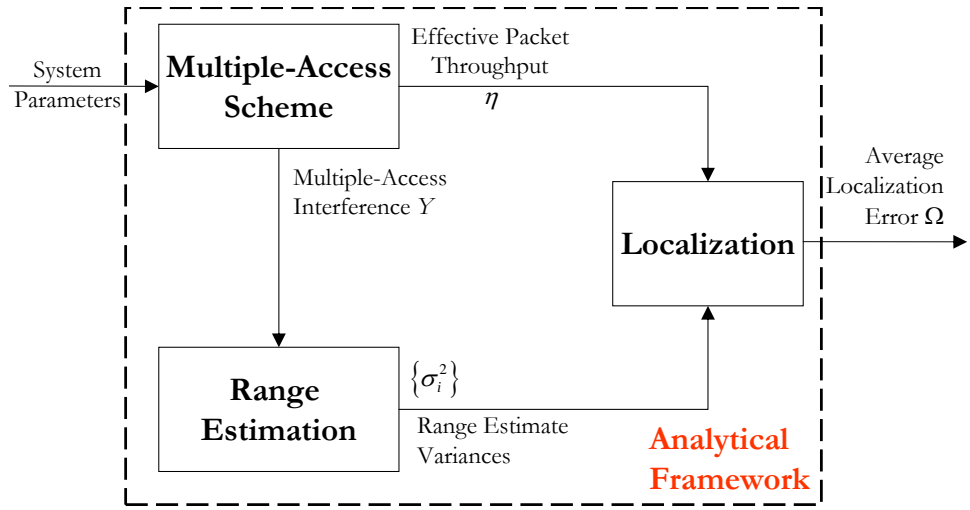


Figure 8.2: Impact of PoLoNet system parameters on the average localization error when contention-based multiple-access schemes are used.

In this case, maximizing the throughput with respect to the given system parameter x does not necessarily minimize the average localization error with respect to x, due to the dependence of the range estimate variance on x. Ideally, we would like to minimize the accuracy of range estimates, while simultaneously maximizing the effective throughput, but varying a given system parameter may have different effects on the accuracy of range estimation and the effective throughput. Thus, we need an analytical framework for modeling the simultaneous impact of network parameters on the effective throughput and the variance of range estimation. As shown in Figure 8.2, this allows us to directly relate the network parameters with the average localization accuracy of an unlocalized node as a function of time. The development of such a framework is discussed in the following sections.

In this chapter, two specific objectives of the developed framework are:

- We would like to analyze the impact of increasing the total number of localized nodes on the average localization accuracy achieved. We saw in Chapter 3 that increasing the number of anchors that provide range estimates improves the average localization accuracy, and the benefits of increasing the number of localized nodes were observed in Chapters 3-6. However, we did not consider the multiple-access design aspects of successfully obtaining range estimates from a large number of localized nodes. This also provides insight into the modeling of more dynamic scenarios: e.g., from the perspective of average localization accuracy, should unlocalized nodes that are able to estimate their locations be allowed to become anchors and provide range estimates to other unlocalized nodes?

- It is also of interest to examine the impact of mobile speed on the average mobile local-

ization accuracy that can be achieved. It is intuitive that as the mobile speed increases, the duration over which range estimates should be gathered and location estimates computed diminishes. The analysis of this tradeoff between localization accuracy and mobile speed provides insight into the selection of network design parameters for a target localization accuracy and mobile speed.

In the following section, we describe the network model used in the development of the discussed analytical framework.

8.3 Network Model

The following assumptions are made in order to associate the average localization accuracy of an unlocalized node with the relevant multiple-access parameters. The multiple-access schemes under consideration are similar to the pure-Aloha [80] scheme adapted to meet the requirements of a MAC protocol for PoLoNets. The assumptions for synchronous and asynchronous PoLoNets are treated separately where relevant. We return to the validity of some of these assumptions in Section 8.6.

Spatial Distribution of Localized and Unlocalized nodes: We assume that the localized and unlocalized nodes are Poisson distributed over the two-dimensional plane, with average spatial densities Λ_L and Λ_U respectively. This implies that the probabilities of finding k localized and l unlocalized nodes in a region with an area A are respectively given by:

$$\begin{aligned} \Pr\{k \text{ localized nodes in an area } A\} &= e^{-\Lambda_L A}\frac{(\Lambda_L A)^k}{k!}, \\ \Pr\{l \text{ unlocalized nodes in an area } A\} &= e^{-\Lambda_U A}\frac{(\Lambda_U A)^l}{l!}. \end{aligned} \tag{8.4}$$

In this case, the area of interest is the entire x-y plane. An important and useful property of a Poisson process [83] is statistical isotropy. This implies that the statistics of the two-dimensional point process are identical relative to any point in the two-dimensional plane. Thus, it suffices to consider the behavior of a single unlocalized node, which, without loss of generality, is located at the origin.

Packet Transmissions in a Synchronous PoLoNet: We assume that the time-axis is slotted, and all nodes are perfectly synchronized to the slot transitions. A single packet can be transmitted within a slot. In such a PoLoNet, the anchor nodes are "packet-sources" and the unlocalized nodes are "packet-sinks". Each anchor node broadcasts a packet in a slot with probability p. The multiple-access scheme is essentially a pure-Aloha scheme [80]

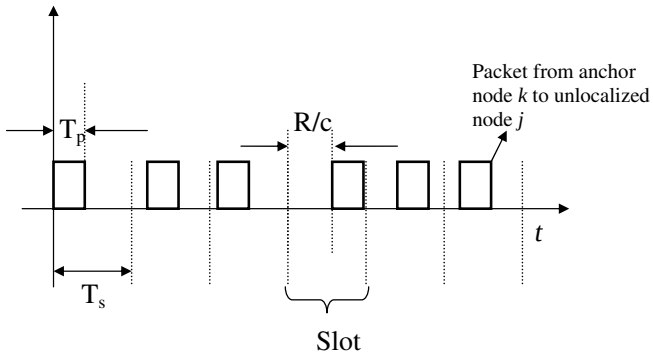


Figure 8.3: Slotted packet transmissions: based on the delay between a packet's arrival-time and the beginning of the slot, an unlocalized node can estimate the distances to transmitting anchors.

with transmission probability p. An unlocalized node that receives a packet from an anchor can estimate the distance R between them based on the TOA of the packet within the slot, as shown in Figure 8.3. Anchor nodes are assumed to transmit packets at slot-transitions, and therefore, the TOA of the packet at an unlocalized node relative to the slot-transition is proportional to R. In order to eliminate range ambiguity, we assume $R \ll cT_s$, where c is the speed of light. Since our goal is to model localization accuracy, the contents of the packets are assumed to be the coordinates of the corresponding anchors. From a communications perspective, these packets could contain additional data, since only the TOAs of the packets are used to obtain range estimates. The transmit power of all localized nodes is assumed to be constant and equal to P_T.

Packet Transmissions in an Asynchronous PoLoNet In this case, we assume that the time-axis is slotted and that although the nodes are not perfectly synchronized to the slot transitions, each node can transmit a single packet within a slot. The slot width T_s is assumed to be greater than the packet duration T_p to accommodate the lack of synchronization between nodes. Due to the lack of synchronization between nodes, range estimation is performed through a packet exchange, as described in Section 2.3. Each unlocalized node broadcasts a RI packet within a given slot indexed i with probability p. Each localized node that successfully decodes this packet, transmits a RR packet *specifically addressed to the unlocalized node that transmitted the RI packet* at the suitable time in time slot $(i+1)$ with probability q. A probability $q < 1$ is used to reduce the probability of collisions of RR packets intended for an unlocalized node that broadcasts a RI packet. If a RR packet can be successfully decoded by the unlocalized node, then a range estimate can be extracted. For simplicity, we assume that RI and RR packets are identical. The above multiple-access

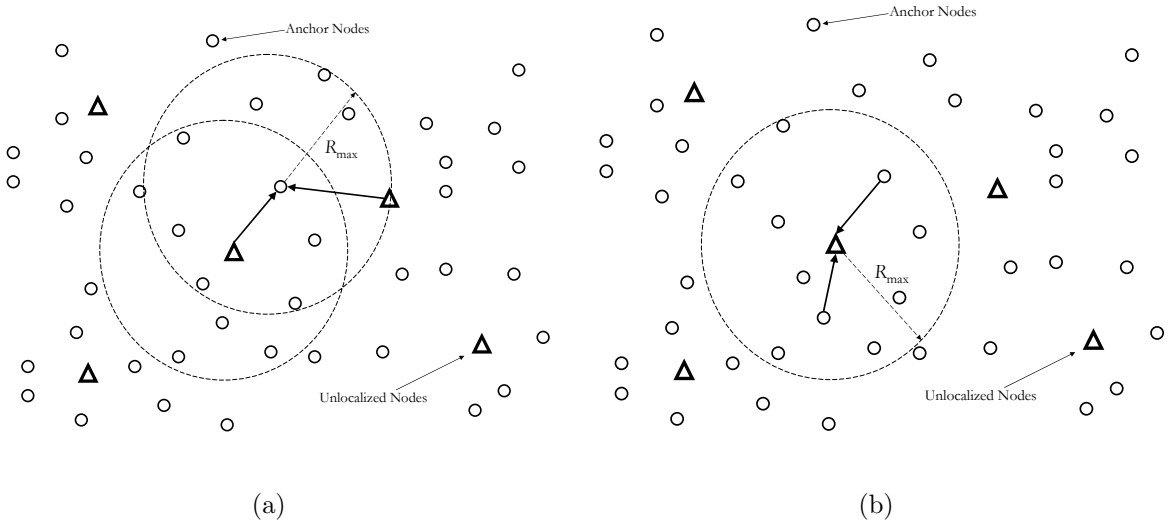


Figure 8.4: Transmissions from (a) unlocalized nodes to localized nodes (RI packets), and (b) localized nodes to unlocalized nodes (RR packets).

scheme can be viewed as a two-way conditional[2] pure-Aloha scheme with transmit probability p for RI packets, and probability q for RR packets.

Figures 8.4(a) and 8.4(b) illustrate the differences in the discussed modeling of synchronous and asynchronous PoLoNets. In synchronous PoLoNets, we restrict our attention to the transmissions from localized to unlocalized nodes shown in Figure 8.4(b). However, in asynchronous PoLoNets, the modeling of the RI transmissions from unlocalized nodes to localized nodes (Figure 8.4(a)), as well as the RR transmissions from localized nodes to unlocalized nodes (Figure 8.4(b)) need to be considered.

Multiple-access interference and the successful reception of packets As localized (and unlocalized nodes in the asynchronous case) can transmit packets simultaneously in a slot, this can result in multiple-access interference (MAI) between simultaneously transmitted packets. We assume that the multiple-access interference seen at an unlocalized node is independent from slot to slot, and further, that the interference power is constant over the length of a slot.

The computation of the statistics of the multiple-access interference presented here closely follows the analysis presented in [31]. We consider the link between an anchor node A and an unlocalized node B. The analysis is analogous for transmissions from B to A, given the nature of the network model assumed. Let the distance between A and B be R. In a given

[2]The transmission of a RR packet is conditioned on the successful reception of an RI packet.

slot, if node A transmits a packet to node B, the received signal power [1] can be modeled as

$$P_{r0} = \frac{K_P P_T}{R^\beta}, \tag{8.5}$$

At B, the received signal power from other interfering anchors that transmit packets in the same slot is given by

$$P_{rk} = \frac{K_P P_T}{R_k^\beta}, \ k = 1, 2, \cdots \quad , \tag{8.6}$$

where R_k is the distance between B and the kth transmitting localized node. The Gaussian model for the MAI in TH-PPM UWB systems was analyzed in [35]. Applying this model to the interference seen at B from other interfering anchors, the Signal-to-Interference-and-Noise Ratio (SINR) S at B can be expressed as

$$S = \left(\frac{1}{\xi_0} + \frac{K_1}{N_s P_{r0}} \sum_k P_{rk} \right)^{-1},$$

where K_1 is a constant dependent on the receiver structure [35], N_s is the spreading gain, and ξ_0 is the Signal-to-Noise ratio (SNR). As the transmit power of nodes is assumed to be identical, using (8.5) and (8.6), this can be rewritten as

$$S = \left(\frac{1}{\xi_0} + \frac{K_1 R^\beta}{N_s} \sum_k \frac{1}{R_k^\beta} \right)^{-1}. \tag{8.7}$$

The SINR S is a random variable, since it depends on the spatial locations of the interfering nodes. Suppose we define the random variable Y as:

$$Y \triangleq \sum_k \frac{1}{R_k^\beta}. \tag{8.8}$$

Then, Y represents the spatial dependence of the total interference power seen from all transmitting localized nodes in a given slot. If Λ' is the average spatial density of transmitting nodes, it can be shown that (see Appendix 8B, [31]) for $\beta > 2$, the characteristic function of Y is given by:

$$\phi_Y(\omega) = \exp\left(-\pi \Lambda' e^{-\frac{\pi}{\beta}} \Gamma\left(1 - \frac{2}{\beta}\right) \omega^{\frac{2}{\beta}} \right).$$

This has the form of the characteristic function of a symmetric α-stable ($S\alpha S$) distribution [84] with a dispersion factor $\alpha = \frac{2}{\beta}$. A closed-form expression for the probability density function and the cumulative distribution function associated with the above characteristic function [85] is known only when $\beta = 4$ (Levy distribution):

$$f_Y(y) \ = \ \frac{\pi \Lambda'}{2} y^{-\frac{3}{2}} e^{-\frac{\pi^3 \Lambda'^2}{4y}}, \ F_Y(y) = \text{erfc}\left(\frac{\pi^{\frac{3}{2}} \Lambda'}{2\sqrt{y}} \right).$$

The expression for the statistical distribution of Y for $\beta > 2, \beta \neq 4$ is known but not in closed-form, and this makes the analysis considerably tedious. We therefore restrict our analysis to the case $\beta = 4$, which is known to be applicable for wireless signal propagation with ground-deployed nodes [31]. We use simulation results in Section 8.6 to verify that similar trends are valid for $\beta = 2$. Using the above statistical distribution for Y, the distribution of the SINR S can be derived using (8.7) and (8.8). This derivation is provided in Appendix 8C.

In the synchronous case, the probability that an anchor node is transmitting in a given slot is p, the set of "active" interfering anchors form a spatial Poisson process [31] with average density $\Lambda' = p\Lambda_L$. In the asynchronous case, we assume that the active interferer density is $\Lambda' = p\Lambda_U + q\Lambda_L$. This is a conservative assumption[3], as the probability of RR transmission is strictly less that q, since it is conditioned on a RI packet being decoded successfully.

A packet can be decoded successfully at a node if the received signal-to-interference-and-noise ratio (SINR) S exceeds a threshold ξ_T, determined by the sensitivity of the receiver and the strength of coding scheme used. We further assume that if the packet is decoded successfully, a range estimate can be obtained through the TOA of the packet within the slot. The maximum distance between two nodes such that a packet can be successfully decoded is denoted by $R_{\max}$. For a spatial Poisson process, the inter-node distance R can be modeled as a uniform random variable:

$$f_R(r) = \begin{cases} \frac{2r}{R_{\max}^2} & r \leq R_{\max}, \\ 0 & \text{otherwise.} \end{cases}$$

Based on the above assumptions, the probability $P_s(R)$ that a packet is decoded successfully can be written as a function of R. As shown in Appendix 8C, when $\beta = 4$, the expression for $P_s(R)$ becomes

$$P_s(R) = \operatorname{erfc}\left(\frac{\pi^{\frac{3}{2}}\Lambda' R^2}{2\sqrt{\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}}\right),$$

where $K_1' = \frac{K_1}{N_s}$. Averaging this probability over the link distance R (between 0 and $R_{\max}$), we obtain

$$\overline{P_s} = \operatorname{erfc}\left(\frac{z}{2}\right) + 2\left(\frac{1 - e^{-\frac{z^2}{4}}}{\sqrt{\pi} z}\right), \; z = \left(\frac{\pi\Lambda'\Gamma(1-\alpha)R_{\max}^2}{\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^{\alpha}}\right). \tag{8.9}$$

Collision model The multiple-access interference (or "secondary interference") is modeled through the statistics of the SINR and $\overline{P_s}$ above. When several packet transmissions with a

[3]This can be replaced by $\Lambda' = p\Lambda_U + qP_{s,RI}\Lambda_L$, where $0 \leq P_{s,RI} \leq 1$ is the probability that a localized node successfully decodes an RI packet transmitted by an unlocalized node in any slot. However, $P_{s,RI}$ is a quantity that we are trying to ascertain through the following analysis, and therefore, we opt for the simpler model.

SINR larger than ξ_T are received at a node, these "primary interferers" need to be regarded differently. In the case where there are j primary interferers, the collision model assumed is that out of the j interfering packets, a packet is randomly selected and decoded with probability $(1-Q)^j$, where Q is termed the capture ratio. The probability that a particular packet out of the j packets is successfully decoded is then given by $\frac{(1-Q)^j}{j}$. This model provides considerable flexibility in terms of the actual multiple-access scheme employed. For instance, in the presence of strong transmission on different spreading codes, we may assume $Q = 0$, and for strong transmissions on the same spreading code, we may assume $Q > 0$. In the following development, we ignore the details of the spreading-code assignment, and assume a constant spreading gain and $Q > 0$ for all transmissions.

TOA-based Range Estimation at unlocalized nodes An unlocalized node that successfully decodes a packet from a localized node can estimate the distance R between them based on the TOA of the packet within the slot and the transmission time of the original RI packet. We assume that the variance of the range estimate received is given by:

$$\sigma^2(R,Y) = \frac{K_R}{S} = K_R\left(\frac{1}{\xi_0} + \frac{K_1 R^\beta Y}{N_s}\right).$$

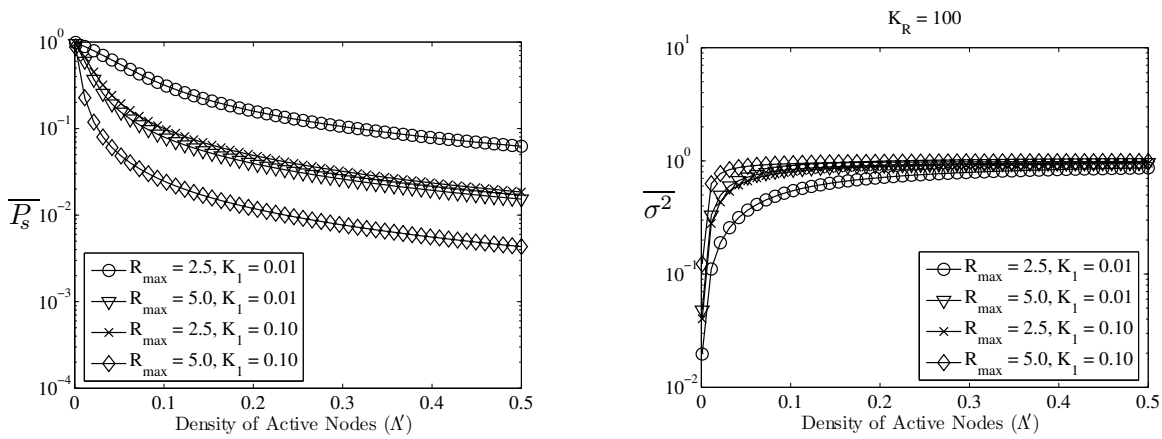
Based on the statistical distribution of Y, we can derive an expression for the variance of range estimates averaged over the statistics of the interference when $\beta = 4$:

$$\sigma^2(R) = \frac{K_R}{P_s(R)}\left[\left(\frac{x\left(\frac{1}{\xi_T}-\frac{1}{\xi_0}\right)}{2}\right)\left(\frac{\exp(-x)}{\sqrt{\pi x}} - \mathrm{erfc}(\sqrt{x})\right) + \frac{1}{\xi_0}\mathrm{erfc}(\sqrt{x})\right], \tag{8.10}$$

where $x = \frac{\pi^3 \Lambda'^2 R^4 K_1'}{\left(\frac{1}{\xi_T}-\frac{1}{\xi_0}\right)}$. The proof of the above equation is given in Appendix 8C. The range estimate variance averaged over the link distance can be obtained using the following expression via numerical integration:

$$\overline{\sigma^2} = \int_R \sigma^2(r) f_R(r) dr = \int_0^{R_{max}} \sigma^2(r)\frac{2r}{R_{\max}^2} dr. \tag{8.11}$$

In (8.9), (8.10) and (8.11) we have expressions for (i) the average probability $\overline{P_s}$ that a packet for a node within a radius $R_{\max}$ can be successfully decoded in the absence of other primary interferers, and (ii) the average variance $\overline{\sigma^2}$ of range estimates that can be estimated based on the TOA of the received packet in the presence of multiple-access interference caused by a Poisson field of interferers with average spatial density Λ'. We assume that in the absence of other primary interferers, for a transmission from *any* node within a distance $R_{\max}$, the probability of packet success is given by $\overline{P_s}$, and the resulting range estimate variance is given by $\overline{\sigma^2}$. This is an oversimplification, as the probability of successful packet transmissions from nearby nodes is expected to be larger than those from more distant nodes. Similarly, we



(a) (b)

Figure 8.5: The average probability $\overline{P_s}$ and the average range error variance $\overline{\sigma^2}$ versus the effective density of nodes Λ'. The value of K_R used is 100. The receiver sensitivity is captured by the parameter $K_1' = \frac{K_1}{N_s}$. The smaller the value of K_1', the more robust the receiver is to multiple-access interference.

expect that distances estimated using packets from nearby nodes should be more accurate than packets from distant nodes. However, including the dependence on distance makes the modeling extremely complicated.

Figure 8.5(a) shows the variation of $\overline{P_s}$ with the effective density Λ'. We see that as Λ' increases, $\overline{P_s}$ decreases and $\overline{\sigma^2}$ increases due to the increase in the interference. As the maximum link distance R_{max} increases, we see that $\overline{P_s}$ decreases due to the decrease in the SINR. Figure 8.5(b) shows the variation of $\overline{\sigma^2}$ with Λ'. The range estimate variance $\overline{\sigma^2}$ tends to saturate as Λ' increases, as the probability of successfully receiving a range estimates falls sharply with increase in Λ'.

Based on the expressions for $\overline{P_s}$ and $\overline{\sigma^2}$ presented above, we derive expressions for the average throughput η of range estimates for synchronous and asynchronous PoLoNets in the following sections. The average localization error can then be computed from η through (8.1). Some identities used in the following development are listed below.

$$\sum_{i=1}^{N} \binom{N}{i} a^i (1-a)^{N-i} b^i = (1-a)^N \sum_{i=1}^{N} \binom{N}{i} \left(\frac{ab}{1-a}\right)^i$$
$$= (1-a)^N \left[\left(1+\frac{ab}{1-a}\right)^N - 1\right] = \left[(1-a+ab)^N - (1-a)^N\right]. \qquad (8.12)$$

This further implies that

$$\sum_{i=1}^{N} \binom{N}{i} a^i (1-a)^{N-i} \left[(1-b_1)^i - (1-b_2)^i\right] = (1-ab_1)^N - (1-ab_2)^N . \quad (8.13)$$

The second identity that is used is:

$$e^{-\alpha} \sum_{n=1}^{\infty} \frac{\alpha^n}{n!} (a^n - b^n) = e^{-\alpha} \left(e^{\alpha a} - 1 - e^{\alpha b} + 1\right) = \left(e^{\alpha(a-1)} - e^{\alpha(b-1)}\right). \quad (8.14)$$

Finally,

$$\sum_{i=0}^{N} \binom{N-1}{i} \frac{a^i}{i+1} = \frac{1}{Na} \left[(1+a)^N - 1\right]. \quad (8.15)$$

8.4 Localization Accuracy in a Synchronous PoLoNet

Without loss of generality, we can restrict our attention to an unlocalized node located at the origin. For the synchronous case, the spatial density of active anchors is given by $\Lambda' = p\Lambda_L$. The expression for $\overline{P_s}$ in this case can be obtained by plugging in $\Lambda' = p\Lambda_L$ into (8.9). Based on the resulting expression for $\overline{P_s}$, the average effective throughput of packets per slot can be derived as follows.

The probability of n transmitters located within a radius $R_{\max}$ of an unlocalized node is:

$$\Pr\{n \text{ transmitters within } R_{\max}\} = e^{-\Lambda'\pi R_{\max}^2} \frac{(\Lambda'\pi R_{\max}^2)^n}{n!}.$$

Based on the discussed model, the probability that k of the n transmission can potentially be decoded successfully is given by $\binom{n}{k}\overline{P_s}^k \left(1-\overline{P_s}\right)^{n-k}$. Given that k transmissions can be received successfully, then the probability that one of these transmissions is received successfully is $(1-Q)^k$. Then, based on the above expressions and (8.14), the effective throughput of range estimates per slot is given by:

$$\begin{aligned}
\eta &= \sum_{n=1}^{\infty} e^{-\Lambda'\pi R_{\max}^2} \frac{(\Lambda'\pi R_{\max}^2)^n}{n!} \sum_{k=1}^{n} \binom{n}{k} \overline{P_s}^k \left(1-\overline{P_s}\right)^{n-k} (1-Q)^k \\
&= e^{-\Lambda'\pi R_{\max}^2} \sum_{n=1}^{\infty} \frac{(\Lambda'\pi R_{\max}^2)^n}{n!} \left[\left(1-Q\overline{P_s}\right)^n - \left(1-\overline{P_s}\right)^n\right] \\
&= e^{-\Lambda'\pi R_{\max}^2 Q\overline{P_s}} - e^{-\Lambda'\pi R_{\max}^2 \overline{P_s}}. \quad (8.16)
\end{aligned}$$

From (8.16), (8.3) and (8.2), an expression for the average localization error in the nth time-slot can be obtained. The expression for $\overline{\sigma^2}$ is obtained from (8.11) by substituting $\Lambda' = p\Lambda_L$.

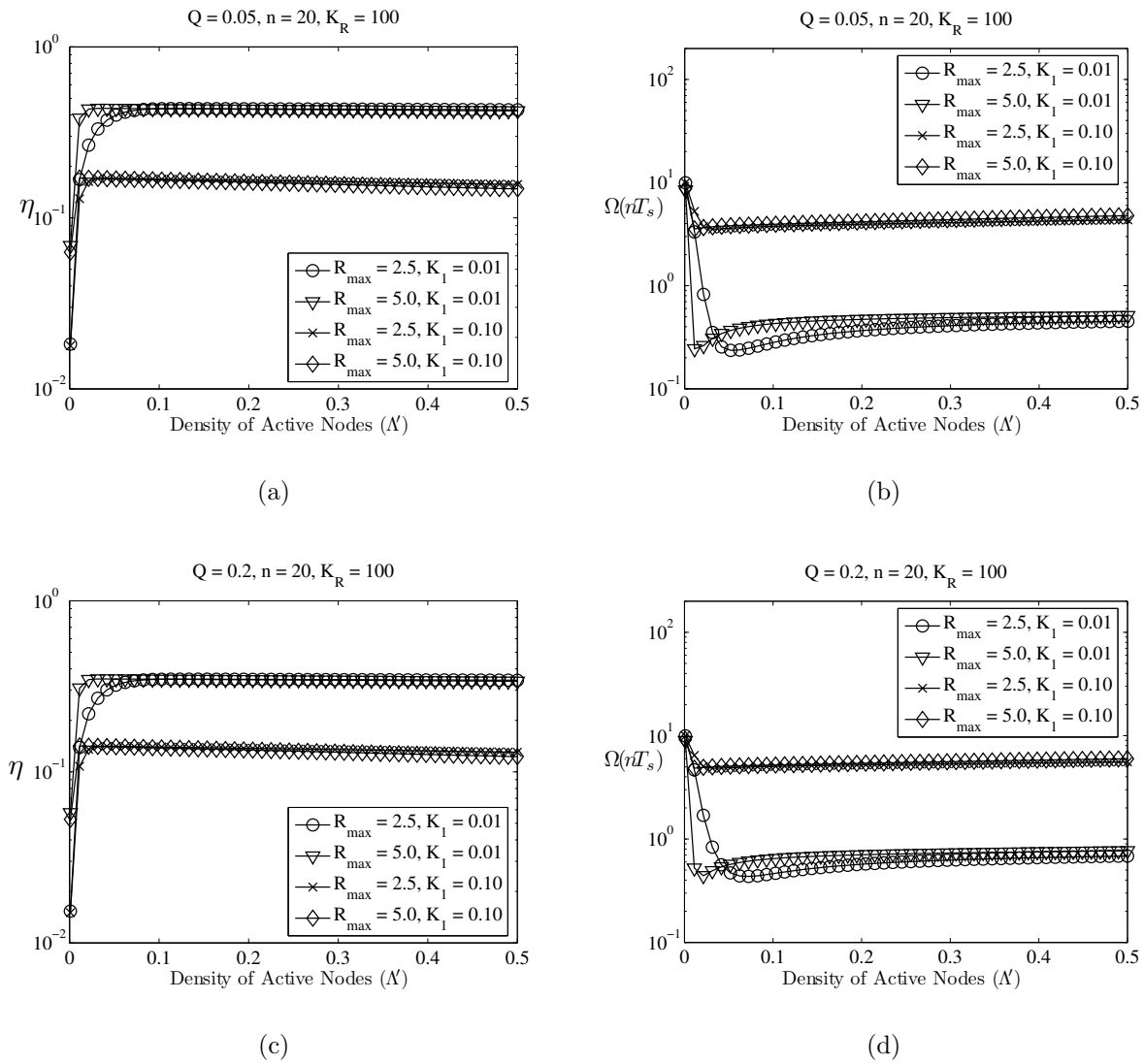


Figure 8.6: The throughput η (Figures (a) and (c)) and the average localization error $\overline{\Omega(n)}$ (meter^2) (Figures (b) and (d)) versus the density of active transmitters Λ' for different values of the capture ratio Q. The values of the other parameters are $n = 20$, $\beta = 4$, $K_R = 100$.

Figure 8.6 shows the variation of the throughput η and the average localization error $\overline{\Omega(n)}$ versus the active transmitter density $\Lambda' = p\Lambda_L$, when $n = 20$. We make the following key observations:

1. There is an optimal value of Λ' at which η is maximized. An explanation for this is as follows: for small values of Λ' there is a small number of anchors within $R_{\max}$ that transmit packets which can be successfully decoded, and as a consequence, η takes on small values. For large values of Λ', η gradually decreases with increase in Λ' due to the increase in multiple-access interference. With regard to the density of localized nodes, this implies that for a fixed p, there exists a Λ_L^* that maximizes η. Analogously, for a fixed Λ_L, there exists a p^* that maximizes η. In order to find the value of a parameter z than maximizes η while the other parameters are kept fixed, we can differentiate η with respect to a generic system parameter z and equate to zero:

$$\begin{aligned} \frac{\partial \eta}{\partial z} = 0 \Rightarrow \frac{\partial}{\partial z} e^{-p\Lambda_L \pi R_{\max}^2 \overline{P_s} Q} - \frac{\partial}{\partial z} e^{-p\Lambda_L \pi R_{\max}^2 \overline{P_s}} = 0, \\ \Rightarrow \quad \left[-Q e^{-p\Lambda_L \pi R_{\max}^2 \overline{P_s} Q} + e^{-p\Lambda_L \pi R_{\max}^2 \overline{P_s}} \right] \left(\frac{\partial p\Lambda_L \pi R_{\max}^2 \overline{P_s}}{\partial z} \right) = 0. \end{aligned} \tag{8.17}$$

   Solving the above equation provides the optimal value of the parameter $z^\star$ that maximizes the throughput.

2. The value of Λ' that maximizes η also minimizes $\overline{\Omega(n)}$. This implies that the value of a parameter that maximizes the average localization accuracy can be obtained by solving (8.17). We further note that although the system parameters impact both the throughput and the variance of range estimates, the average localization accuracy is maximized when the throughput is maximized.

3. As the receiver sensitivity and the robustness to secondary multiple-access interference (parametrized by K_1') increases, the throughput and the localization accuracy increase. We also see that as Q increases, the effective throughput of packets decreases.

In the following section, we extend the above analysis to the case of asynchronous PoLoNets.

8.5 Localization Accuracy in a Asynchronous PoLoNet

In this case, the spatial density of active anchors that generate secondary interference is modeled by $\Lambda' = p\Lambda_U + q\Lambda_L$. The expression for $\overline{P_s}$ in this case can be obtained by plugging in $\Lambda' = p\Lambda_U + q\Lambda_L$ into (8.9). Based on the resulting expression for $\overline{P_s}$, the average effective throughput of packets per slot can be derived as follows. As before, consider an unlocalized

node B located at the origin. Node B broadcasts an RI packet in a given slot with probability p. The probability that there are n localized nodes within a distance $R_{\max}$ of B is:

$$\Pr\left\{n \text{ localized nodes within a distance } R_{\max}\right\} = e^{-\Lambda_L \pi R_{\max}^2} \frac{\left(\Lambda_L \pi R_{\max}^2\right)^n}{n!}. \tag{8.18}$$

The probability that j of the n localized nodes in the neighborhood of B receive the RI packet successfully is given by:

$$\Pr\left\{j \text{ of } n \text{ localized nodes receive RI packet}\right\} = \binom{n}{j} \overline{P_{s,RI}}^j \left(1 - \overline{P_{s,RI}}\right)^{n-j}, \tag{8.19}$$

where $P_{s,RI}$ is the probability that an RI packet is received successfully by a localized node within a distance $R_{\max}$, which will be derived shortly. Since localized nodes that successfully receive an RI packet only respond with probability q, the probability that k of the j localized nodes that received the RI packet from B respond by transmitting an RR packet to B in the following slot is

$$\Pr\left\{k \text{ of } j \text{ localized nodes transmit RR packet}\right\} = \binom{j}{k} q^k \left(1-q\right)^{j-k}. \tag{8.20}$$

Finally, the probability that one of the k RR packets is received successfully is given by:

$$\begin{aligned}\Pr\left\{\text{RR packet received successfully}\right\} &= \sum_{l=1}^{k} \binom{k}{l} (1-Q)^l \overline{P_s}^l \left(1 - \overline{P_s}\right)^{k-l} \\ &= \left[\left(1 - \overline{P_s} Q\right)^k - \left(1 - \overline{P_s}\right)^k\right]. \end{aligned} \tag{8.21}$$

using the identity (8.12). Therefore, the average probability that node B receives an RR packet in the following slot after transmitting an RI packet is given by:

$$\begin{aligned}\overline{P_{s,RR}} &= \sum_{n=1}^{\infty} e^{-\Lambda_L \pi R_{\max}^2} \frac{\left(\Lambda_L \pi R_{\max}^2\right)^n}{n!} \sum_{j=1}^{n} \binom{n}{j} \overline{P_{s,RI}}^j \left(1 - \overline{P_{s,RI}}\right)^{n-j} \times \\ &\qquad \times \sum_{k=1}^{j} \binom{j}{k} q^k \left(1-q\right)^{j-k} \left[\left(1 - \overline{P_s} Q\right)^k - \left(1 - \overline{P_s}\right)^k\right] \\ &= \sum_{n=1}^{\infty} e^{-\Lambda_L \pi R_{\max}^2} \frac{\left(\Lambda_L \pi R_{\max}^2\right)^n}{n!} \sum_{j=1}^{n} \binom{n}{j} \overline{P_{s,RI}}^j \left(1 - \overline{P_{s,RI}}\right)^{n-j} \left[\left(1 - q\overline{P_s} Q\right)^j - \left(1 - q\overline{P_s}\right)^j\right] \\ &= \sum_{n=1}^{\infty} e^{-\Lambda_L \pi R_{\max}^2} \frac{\left(\Lambda_L \pi R_{\max}^2\right)^n}{n!} \left[\left(1 - \overline{P_{s,RI}} q \overline{P_s} Q\right)^n - \left(1 - \overline{P_{s,RI}} q \overline{P_s}\right)^n\right] \\ &= e^{-\Lambda_L \pi R_{\max}^2 \overline{P_{s,RI}} q \overline{P_s} Q} - e^{-\Lambda_L \pi R_{\max}^2 \overline{P_{s,RI}} q \overline{P_s}}, \end{aligned}$$

where the identities (8.12) and (8.13) have been used for simplification. Since an unlocalized node transmits an RI packet in an arbitrary slot with probability p, and since the range

estimates are received in the following slot, the normalized throughput of range estimates per slot is given by:

$$\eta = \frac{p}{2}\overline{P_{s,RR}} = \frac{p}{2}\left(e^{-\Lambda_L \pi R_{\max}^2 \overline{P_{s,RI}} q \overline{P_s} Q} - e^{-\Lambda_L \pi R_{\max}^2 \overline{P_{s,RI}} q \overline{P_s}}\right). \tag{8.22}$$

8.5.1 Expression for $\overline{P_{s,RI}}$

Consider a localized node A. Node A receives RI packets from unlocalized nodes within the radius $R_{\max}$. The probability that there are m transmitting unlocalized nodes in the neighborhood of A is:

$$\Pr\{m \text{ transmitting unlocalized nodes within a distance } R_{\max}\} = e^{-p\Lambda_U \pi R_{\max}^2} \frac{(p\Lambda_U \pi R_{\max}^2)^m}{m!}.$$

The probability that the RI packet from node B is received successfully is constituted by the probabilities that (i) the packet from B has a sufficiently high SINR to be successfully decoded, and (ii) the packet from node B is selected from the remaining i, $0 \le i \le m-1$ primary interferers:

$$\begin{aligned}
&\Pr\{\text{RI packet from node } B \text{ is received}\} \\
&= \overline{P_s} \sum_{i=0}^{m-1} \binom{m-1}{i} \overline{P_s}^i \left(1-\overline{P_s}\right)^{m-i-1} (1-Q)^{i+1} \cdot \frac{1}{i+1} \\
&= \overline{P_s}(1-Q)\left(1-\overline{P_s}\right)^{m-1} \sum_{i=0}^{m-1} \binom{m-1}{i} \left(\frac{\overline{P_s}(1-Q)}{1-\overline{P_s}}\right)^i \cdot \frac{1}{i+1} \\
&= \frac{1}{m}\left[\left(1-Q\overline{P_s}\right))^m - \left(1-\overline{P_s}\right)^m\right].
\end{aligned}$$

Therefore,

$$\begin{aligned}
\overline{P_{s,RI}} &= e^{-p\Lambda_U \pi R_{\max}^2} \sum_{m=1}^{\infty} \frac{(p\Lambda_U \pi R_{\max}^2)^m}{m \cdot m!} \left[\left(1-Q\overline{P_s}\right))^m - \left(1-\overline{P_s}\right)^m\right] \\
&= e^{-p\Lambda_U \pi R_{\max}^2} \left[\sum_{m=1}^{\infty} \frac{(p\Lambda_U \pi R_{\max}^2)^m}{m \cdot m!} \left(1-Q\overline{P_s}\right))^m - \sum_{m=1}^{\infty} \frac{(p\Lambda_U \pi R_{\max}^2)^m}{m \cdot m!} \left(1-\overline{P_s}\right)^m\right] \\
&= e^{-p\Lambda_U \pi R_{\max}^2} \left[T\left(p\Lambda_U \pi R_{\max}^2 \left(1-Q\overline{P_s}\right)\right) - T\left(p\Lambda_U \pi R_{\max}^2 \left(1-\overline{P_s}\right)\right)\right],
\end{aligned} \tag{8.23}$$

where

$$T(x) = \sum_{m=1}^{\infty} \frac{x^m}{m \cdot m!}.$$

The function $T(x)$ can be further expressed as:

$$\begin{aligned} T(x) &= \sum_{m=1}^{\infty} \frac{x^m}{m \cdot m!} = \sum_{m=1}^{\infty} \frac{1}{m!} \int_0^x u^{m-1} du = \int_0^x \frac{1}{u} \sum_{m=1}^{\infty} \frac{1}{m!} u^m du \\ &= \int_0^x \left(\frac{e^u - 1}{u} \right) du = Ei(x) - \log x + \gamma, \end{aligned}$$

where $Ei(x)$ is the exponential integral function, and $\gamma = 0.5772$ is Euler's constant. Therefore, the effective throughput of range estimates in the asynchronous case is defined by (8.22) and (8.23). The average range estimate variances can be obtained through (8.11) with $\Lambda' = p\Lambda_U + q\Lambda_L$. As before, from (8.16), (8.3) and (8.2), an expression for the average localization error in the nth time-slot can be obtained.

Figure 8.7 shows the variation of the effective throughput η computed using (8.22), and the corresponding average localization error $\overline{\Omega(n)}$ where $n = 1000$. In these simulations, $p = q = 0.2$, $\beta = 4$, and $K_R = 100$. We see that there is an optimal set of node densities $(\Lambda_U^*, \Lambda_L^*)$ at which η is maximized, due to the interference-limited nature of the system. As observed in the synchronous case, we see that maximizing η results in the minimization of the localization error. We observe that the value of η is much smaller than the throughput obtain in the synchronous case, as obtaining a range estimate in the asynchronous case is contingent on two successful packet receptions, as opposed to one in the synchronous case.

Figure 8.8 shows the variation of η with the probabilities of RI and RR packet transmissions p and q respectively. It was verified in this case as well that the optimal values (p^*, q^*) that maximize the effective throughput also minimize the average localization error for all values of n. Once again, the optimal value of a system parameter z with respect to localization accuracy can be extracted using:

$$\frac{\partial \eta}{\partial z} = 0 \Rightarrow \frac{\partial}{\partial z} \left(p \left(e^{-\Lambda_L \pi R_{\max}^2 \overline{P_{s,RI}} q \overline{P_s} Q} - e^{-\Lambda_L \pi R_{\max}^2 \overline{P_{s,RI}} q \overline{P_s}} \right) \right) = 0. \tag{8.24}$$

8.5.2 Variation of the Localization Error with mobile speed

In general, there is a trade-off between the rate of change of a quantity and the accuracy with which the quantity can be measured. Intuitively, such a trade-off must occur between mobile localization accuracy and mobile speed. As the mobile speed increases, the duration over which mobile location estimates need to be computed decreases. A plausible means of compensating for the increase in mobile speed is to increase the transmission probability p of RI packets. In order to evaluate whether this response to an increase in the mobile speed is indeed valid, we attempt to characterize the relation between the maximum allowable mobile speed at the current localization accuracy can be guaranteed, in terms of the system parameters.

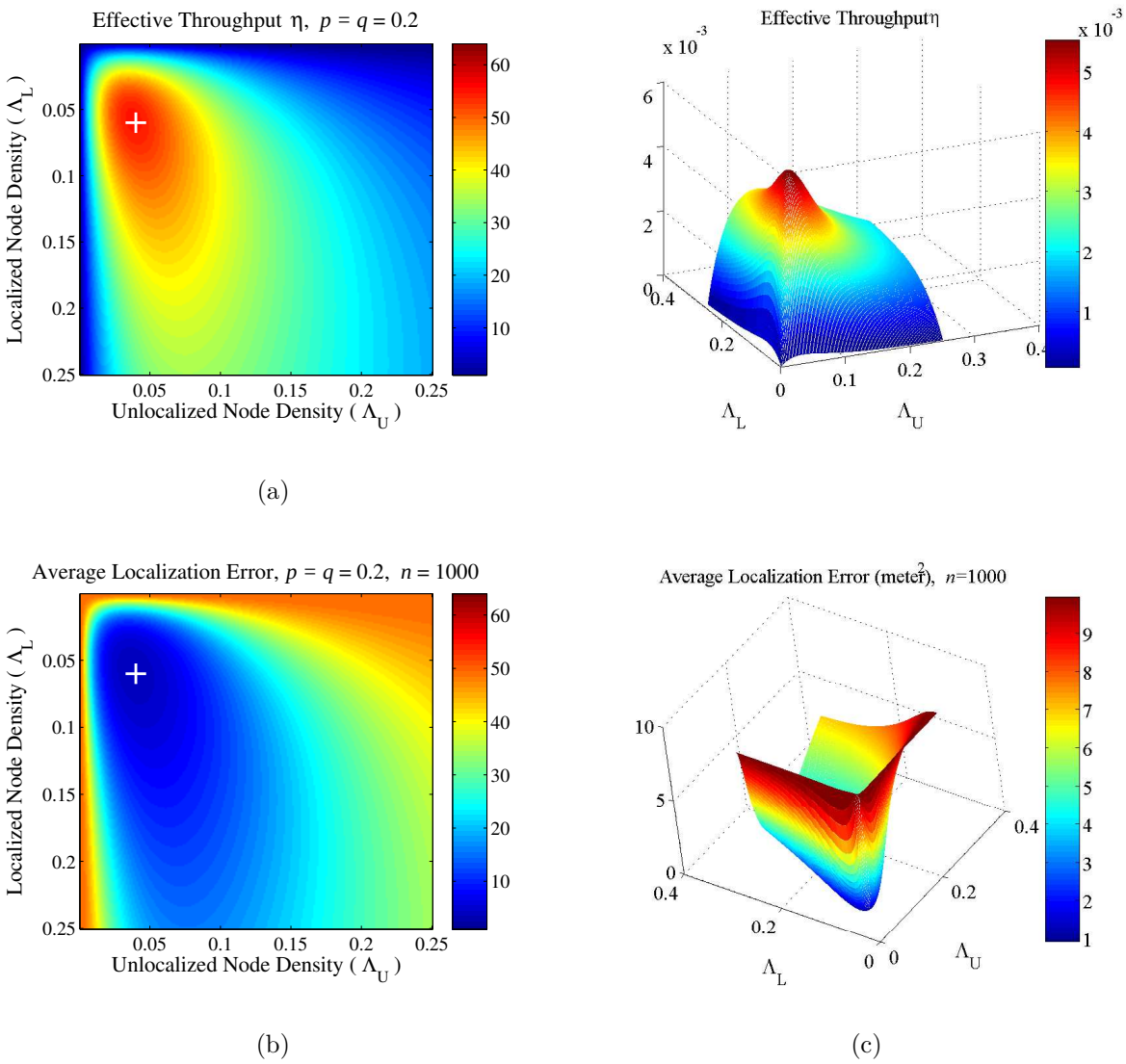


Figure 8.7: Variation of η and $\overline{\Omega(n)}$ with the localized and unlocalized node densities, Λ_L and Λ_U, when $n = 1000$. The values of η and $\overline{\Omega(n)}$ have been scaled in Figures (a) and (c) to ensure maximum contrast. The extreme values are indicated by the white "+" marks. Here, $p = q = 0.2$, $\beta = 4$, $R_{\max} = 5$ meters, $K_1' = 0.01$, and $K_R = 100$.

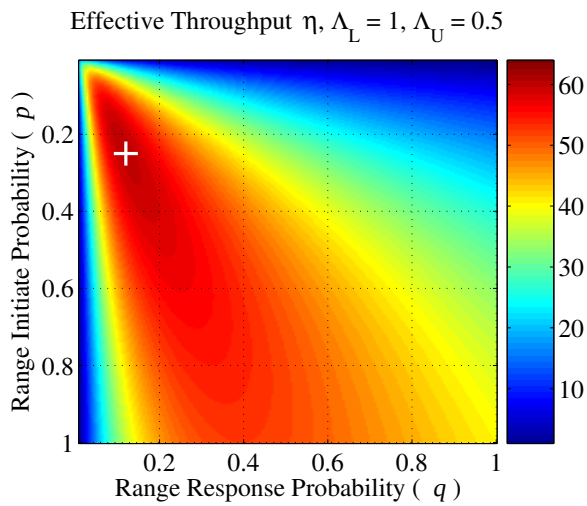


Figure 8.8: Variation of the effective throughput η with the range initiate and range response probabilities p, q. The throughput has been scaled by a factor of 1000 to increase image contrast. The localized and unlocalized node densities are fixed at $\Lambda_L = 1$ nodes/meter^2, and $\Lambda_L = 0.5$ nodes/meter^2. The other parameters have the same values as in Figure 8.7.

In order to derive a framework for the relationship between the maximum mobile speed and the network parameters, we begin by stating the following definition for the maximum allowable mobile speed:

$$v_{\max}\overline{T_{loc}} \triangleq \sqrt{g\overline{\Omega}}, \tag{8.25}$$

where $v_{\max}$ is the maximum mobile speed, $\overline{T_{loc}}$ is the average "localization time", g is a scale parameter, and $\overline{\Omega}$ is the average localization accuracy that is achieved. This relation suggests that if the root localization error caused by the motion of the mobile during the localization time is a factor g larger than the localization accuracy that can be achieved when the node is stationary, then mobility is the dominant source of localization error. Based on this definition, we can relate $v_{\max}$ to the network parameters through the average localization time $\overline{T_{loc}}$ and the average localization error. We further define the localization time as the time take to gather m_0 range estimates [4]. The probability that the localization time $T_{loc} = kT_s$ is given by

$$\Pr\{T_{loc} = kT_s\} = \eta \binom{k-1}{m_0-1} (1-\eta)^{k-m_0} \eta^{m_0-1}. \tag{8.26}$$

This essentially denotes the probability that the m_0th range estimate is received in the nth time slot, and $m_0 - 1$ range estimates are received within the first $n-1$ slots. If $T_{\max}$ is

[4]Strictly speaking, the localization time should be defined as the time take to achieve a certain localization accuracy, but this makes the analysis extremely complex. The main issue with the definition above is that it does not stipulate that range estimates from *different* localized nodes are received.

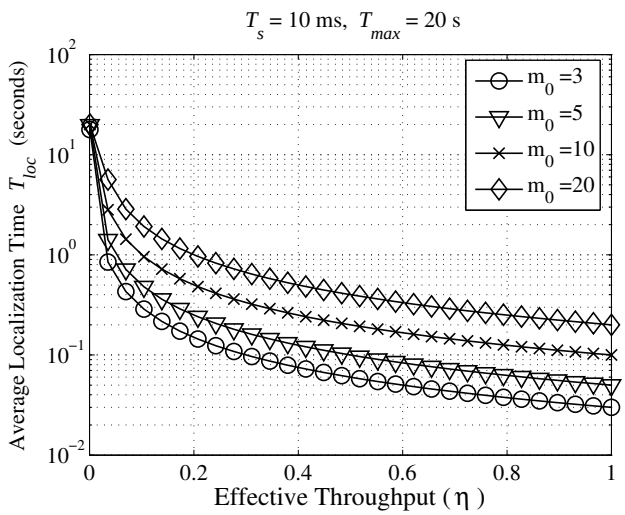


Figure 8.9: Average localization time T_{loc} versus the effective throughput η for different values of m_0. In the above figure, the slot width $T_s = 10$ ms, and the maximum localization time $T_{\max} = 20$ seconds.

the maximum allowable localization time, this occurs with the probability that less than m_0 range estimates were received in k slots is:

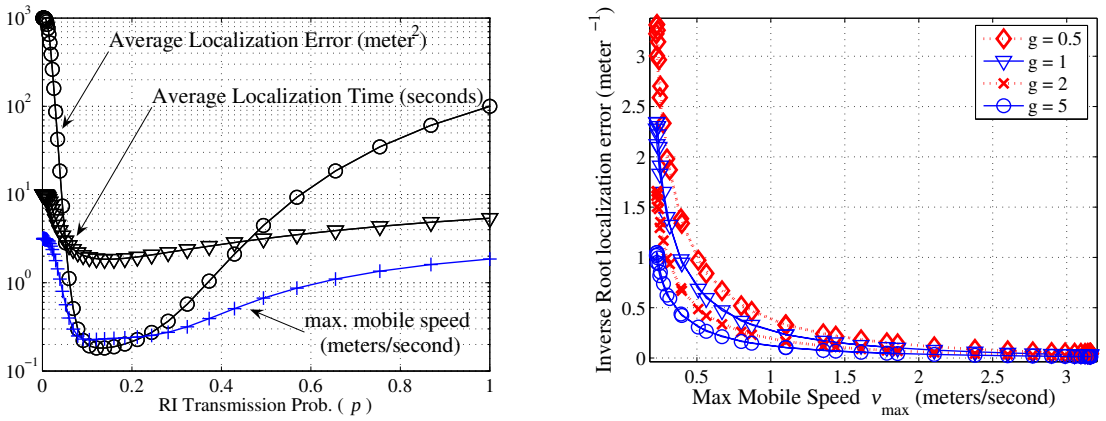
$$\Pr\{\text{Less than } m_0 \text{ range estimates received}\} = \sum_{i=0}^{m_0-1} \binom{m_0-1}{i} \eta^i (1-\eta)^{m_0-i-1}.$$

The average localization time is then given by

$$\begin{aligned} \overline{T_{loc}} &= T_s \sum_{k=m_0}^{\infty} k \Pr\{T_{loc} = kT_s\} \\ &= T_{\max} \sum_{i=0}^{m_0-1} \binom{m_0-1}{i} \eta^i (1-\eta)^{m_0-i-1} + T_s \eta^{m_0} \sum_{k=m_0}^{\infty} k \binom{k-1}{m_0-1} (1-\eta)^{k-m_0}. \end{aligned}$$

Thus, through the definition of η in (8.22), we can express the localization time $\overline{T_{loc}}$ in terms of the system parameters. Subsequently, $v_{\max}$ can be obtained using (8.25) and the localization error defined in (8.3). Figure 8.9 illustrates the variation of $\overline{T_{loc}}$ with η, and as expected, $\overline{T_{loc}}$ decreases as η increases. Evidently, as the number of range estimates m_0 required to form a reliable estimate increases, the average localization time also increases.

Figure 8.10 illustrates the analytical modeling of $v_{\max}$ through the variation of the RI transmission probability p. Figure 8.10(a) illustrates the variation of the average localization error, the average localization time, and the maximum mobile speed with p. The value of g is assumed to be unity. We see that $v_{\max}$ is minimized when the average localization error is



(a) (b)

Figure 8.10: (a) The variation of the average localization error, the average localization time, and the maximum mobile speed ($g = 1$) as p increases, and (b) the reciprocal of the average root localization error (meter^{-1}) versus maximum mobile speed (meters/second) as p is varied for different values of g. The values of the parameters used in this simulation are: $q = 0.2$, $K_R = 100$, $T_s = 10$ ms, $\beta = 4$, $\Lambda_L = 0.5$ nodes/meter^2, $\Lambda_U = 0.25$ nodes/meter^2, $m_0 = 3$, $K_1' = 0.01$, $R_{\max} = 5$ meters, $Q = 0.05$.

minimized. This suggests a tradeoff between localization accuracy and mobile speed, which is illustrated in Figure 8.10(b). Figure 8.10(b) plots the reciprocal of the root localization error versus the maximum mobile speed as p is varied, for different values of g. We see that as the localization accuracy (y-axis) increases, the maximum allowable mobile speed decreases and vice versa, as expected. As g increases, a given localization accuracy can be attained at higher mobile speeds. We note that increasing the value of p as the mobile speed increases does not improve the average localization error.

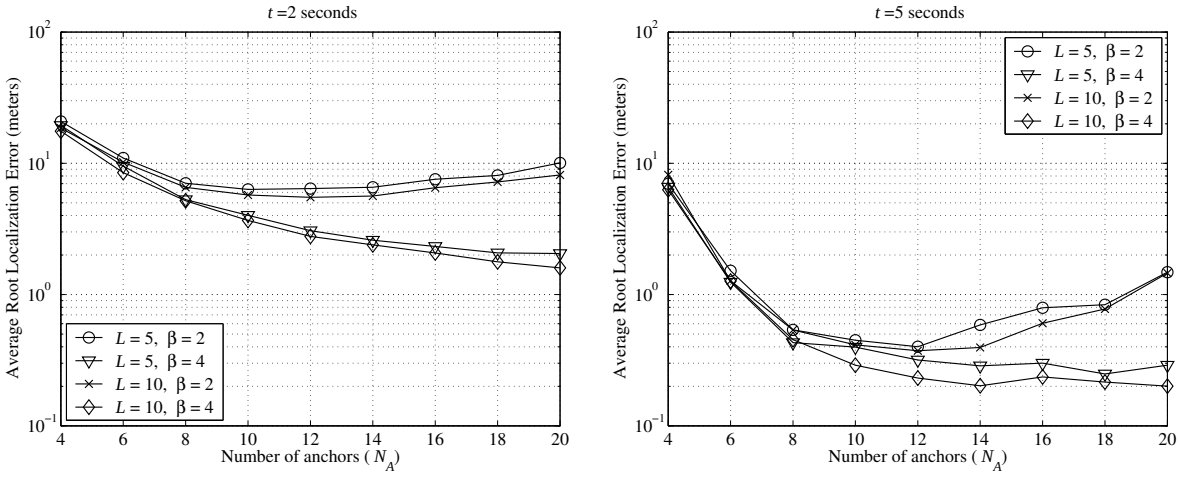
In this section, we have established a framework for the analysis of the interaction between localization accuracy and relevant multiple-access and network parameters. However, the analytical results presented correspond to the case $\beta = 4$, and several simplifying assumptions were required in order to make the analysis tractable. In the following section, we present simulation results that evaluate the validity of the insights gain through the above analytical development.

8.6 Comparison with Simulation Results, Modeling Issues and Further Work

Based on the simulation framework discussed in section 7.4, the performance of the Aloha-based scheme discussed in Section 8.3 was simulated. Figure 8.11 shows the variation of the simulated values of the average root localization error (in meters) of $N_R = 5$ stationary unlocalized nodes randomly located over an $L \times L \times L$ volume of interest. Here, $N_A \geq 4$ anchors are randomly placed within the area of interest. The location estimates are computed using the LS estimator, and all links are assumed to be LOS. Figures 8.11(a) and 8.11(b) indicate the root localization error versus N_A at $t = 2$ seconds and $t = 5$ seconds respectively.

As seen in the analytical results, there is an optimal density of localized nodes at which the localization accuracy is maximized. In this case, when $\beta = 2$, the optimal number of anchors is $N_A^* = 10$, and when $\beta = 4$, $N_A^* > 20$. This is an intuitive result as the average localization accuracy is likely to be more strongly limited by multiple-access interference when $\beta = 2$, as opposed to $\beta = 4$. Therefore, from a multiple-access perspective, increasing the number of anchors beyond a certain value degrades the average localization accuracy due to the increase in multiple-access interference. For small values of N_A, the number of range estimates is too small, when N_A is large, multiple-access interference reduces the throughput of range estimates and their accuracy.

Figure 8.12(a) presents a simulation of the average rate of transition in the density of unlocalized nodes and localized nodes as a function of time. In this simulation, we assume that at $t = 0$, the number of anchors $N_A = 5$, the number of unlocalized reference nodes $N_R = 5$, and $L = 10$ meters. Here, unlocalized nodes that are able to receive $m_0 = 4$ range estimates from *distinct* localized nodes estimate their locations, and then act as anchors for other unlocalized nodes. The time-varying densities of localized and unlocalized nodes are



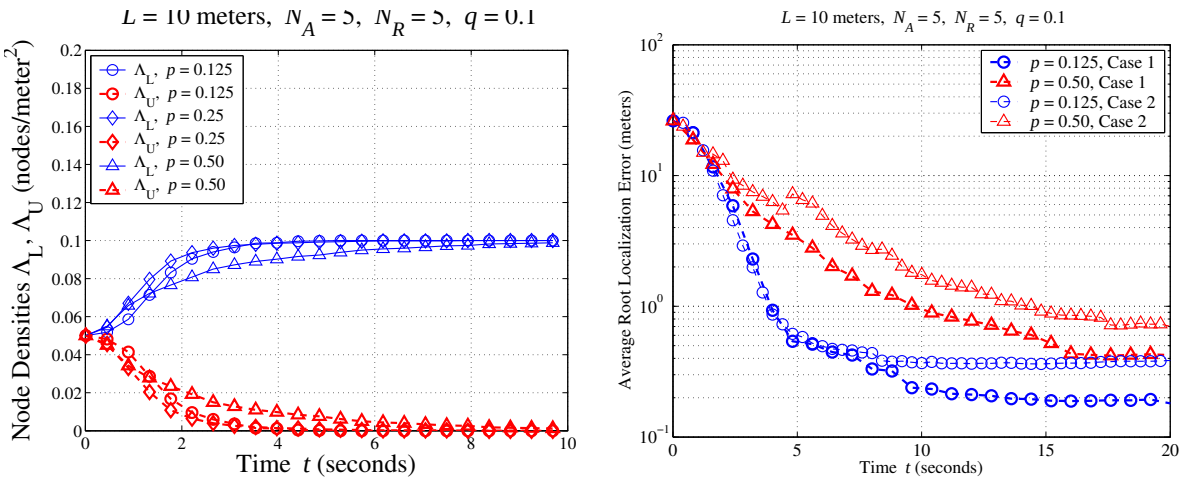
(a) (b)

Figure 8.11: Simulated values of the average root localization error of stationary reference nodes as the number of anchors N_A is varied. The root localization error curves at (a) $t = 2$ seconds, and (b) $t = 5$ seconds are shown for different values of L and β.

computed by dividing the number of localized and unlocalized nodes at any instant by L^2. We see from Figure 8.12(a) that the choice of the parameter p plays a significant role in determining the "network localization time", or the time taken for all unlocalized nodes to become localized. By choosing p suitably, the network localization time can be reduced.

Figure 8.12(b) compares the average root localization error of nodes in two cases: (1) unlocalized nodes that compute their location estimates are not allowed to become anchors, (2) unlocalized nodes that compute their location estimates are allowed to become anchors. The values of the parameters in this simulation were: $\beta = 2$, $L = 10$ meters, $q = 0.1$, $N_A = 5$, and $N_R = 5$. We see that Case 1 can result in considerably improved performance relative to Case 2 for certain parameter choices. This implies that in order to take advantage of the presence of additional localized nodes within the area of interest as unlocalized nodes estimate their locations, the system parameters need to be suitably selected (possibly in a time-variant manner) in order to reduce the increase in multiple-access interference.

However, the extraction of the optimal parameters to be used in this dynamic scenario is not straightforward. In our analytical framework, we assumed that the number of localized and unlocalized nodes were fixed, and that the average throughput of range estimates was constant over all time-slots. We further showed that the throughput of range estimates depends on the densities of localized and unlocalized nodes. However, if the densities of localized and unlocalized nodes are time-varying, this implies that the throughput is also



(a) (b)

Figure 8.12: (a) Change in the densities of localized and unlocalized nodes versus time. Here, $N_A = 5$ anchors are located within the area of interest, which is an $L \times L \times L$ volume, $L = 10$ meters, and $N_R = 5$ is the number of stationary unlocalized nodes. (b) Average root localization error in two cases (i) Case 1: Unlocalized nodes that compute their location estimates are not allowed to become anchors, (ii) Case 2: Unlocalized nodes that compute their location estimates are allowed to become anchors.

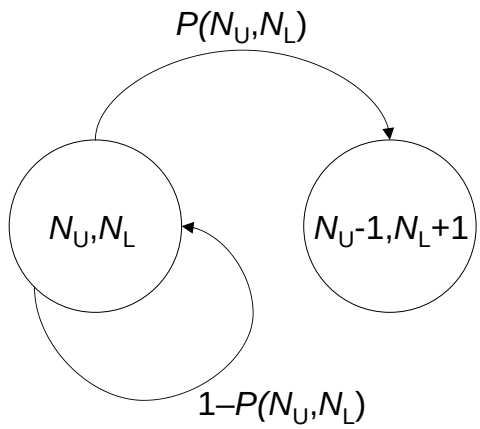


Figure 8.13: Probabilistic state model for node transitions: the probability of an unlocalized node becoming localized in a given slot is a function of the number of localized and unlocalized nodes, and other system parameters.

time-variant, and the corresponding analysis becomes extremely complicated. A possible approach to this problem is to develop a simple probabilistic model for the transition of unlocalized nodes to localized nodes, as shown in Figure 8.13. The transition probabilities at time instant depend on the current values of the system parameters, and the current densities of localized and unlocalized nodes. The development of such a model can provide insights into the behavior of more dynamic scenarios.

Finally, Figure 8.14 presents a simulation of the average localization accuracy (the average inverse root localization error) of a mobile node traveling linearly with a speed v through a network of $N_A = 20$ localized nodes deployed over an $L \times L \times L$ volume, $L = 10$ meters. We see that the trends observed are similar to those developed through the analytical framework. Further, we see that if target localization accuracy are the expected mobile speeds are provided, performance enhancement through the selection of system parameters may be possible.

8.6.1 Modeling Issues

In the modeling of the synchronous and asynchronous PoLoNets, several simplifying assumptions were made in an effort to make the analysis simple yet insightful. However, in order to gain a better understanding of the behavior of practical PoLoNets, further investigation is required into more realistic models for such networks from the perspective of localization accuracy. Below, we list some aspects of the modeling of location-aware sensor networks that warrant investigation and require more sophisticated modeling strategies.

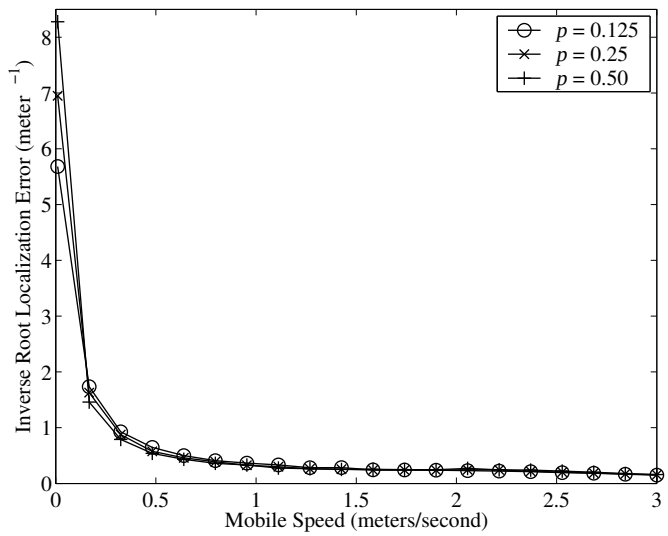


Figure 8.14: Simulated values of the average inverse root localization error (meter^{-1}) of a mobile versus the speed of mobile (meters/second). Here, $\beta = 4$, $N_A = 20$, $L = 10$ meters. The anchor nodes are randomly dispersed over the $L \times L \times L$ volume, and the mobile node travels linearly across the volume of interest with a speed v. The curves plotted correspond to different values of p.

Dependence on the true inter-node distances: In this chapter, we ignored the dependence of the probability of successful packet transmission and the variance of range estimates on the distances between unlocalized and localized nodes. We assumed that the probability of successful packet transmission and the variance of range estimates for all unlocalized and localized nodes within a distance $R_{\max}$ is constant and are given by $\overline{P_s}$ and $\overline{\sigma^2}$ respectively. In reality, the the probability of successful packet transmissions from nearby nodes is expected to be larger than those from more distant nodes. Similarly, we expect that distances estimated using packets from nearby nodes should be more accurate than packets from distant nodes. Therefore, a more accurate model for the localization accuracy would include the dependence of the localization accuracy on the distribution of the inter-node distances.

Poisson Distribution of Nodes The spatial Poisson distribution for the locations of nodes has several advantages and attractive mathematical properties. The statistical isotropy and memoryless properties of the Poisson process make the analysis of the interference from Poisson distributed terminals tractable. However, the main disadvantage of using the spatial Poisson assumption for the locations of nodes is that the area of the network is assumed to be infinite. Models for the behavior of PoLoNets deployed over a finite area need to be developed.

Fairness: The issue of fairness of packet transmissions has not be considered in the analysis of throughput. For instance, it may be possible to maximize the average throughput of range estimates by allowing a single localized node to transmit repeatedly to an unlocalized node, i.e., we do not enforce the condition that range estimates be received from m_0 distinct localized nodes in the analysis. Due to the probabilistic nature of the Aloha-based multiple-access scheme discussed, fairness was not found to be an issue via simulations. However, the interpretation of the general relation between localization accuracy and the throughput necessitates the discussion of fairness.

Slotted versus Unslotted: While discussing the performance of asynchronous PoLoNets we assumed that the time axis was slotted, and it was assumed that each node can transmit a single packet within a slot. Unslotted packet transmissions based on continuous-time Poisson packet-arrivals can be investigated.

Collision model and code-assignment In the development of the analytical framework for the impact of multiple-access schemes on localization accuracy, the specific details of spreading code assignment were not discussed. The capture ratio Q was used model the effect of primary interferers, but specific models to incorporate the details of the code-assignment need to be incorporated.

8.7 Conclusions

In this chapter, we presented an analytical framework for characterizing the dependence of localization accuracy on network parameters such as node densities, mobile speed and other multiple-access parameters in synchronous and asynchronous PoLoNets. The insights into the impact of these parameters on the localization accuracy provided by the developed analytical framework were then verified using simulation results. We found that (i) from a multiple-access perspective, there is an optimal density of localized nodes that maximizes the average localization accuracy, and (ii) there is a trade-off between the mobile speed and the mobile localization accuracy. The insights gained from the analytical framework further provide an approach to modeling the more dynamic nature of practical ad hoc PoLoNets. A discussion of possible improvements to the modeling of PoLoNets from a multiple-access perspective was provided.

8.8 Appendix 8A: On the relationship between localization accuracy and effective throughput

Define $H_{m,n} = \eta^m (1-\eta)^{n-m}$. Then,

$$\frac{\partial}{\partial x} H_{m,n} = \frac{\partial H_{m,n}}{\partial \eta} \cdot \frac{\partial \eta}{\partial x},$$

and

$$\begin{aligned}
\frac{\partial^2}{\partial x^2} H_{m,n} &= \frac{\partial}{\partial x}\left(\frac{\partial H_{m,n}}{\partial \eta} \cdot \frac{\partial \eta}{\partial x}\right), \\
&= \left(\frac{\partial}{\partial x}\frac{\partial H_{m,n}}{\partial \eta}\right) \cdot \frac{\partial \eta}{\partial x} + \frac{\partial H_{m,n}}{\partial \eta} \cdot \frac{\partial^2 \eta}{\partial x^2}, \\
&= \left(\frac{\partial^2 H_{m,n}}{\partial \eta^2}\right) \cdot \left(\frac{\partial \eta}{\partial x}\right)^2 + \frac{\partial H_{m,n}}{\partial \eta} \cdot \frac{\partial^2 \eta}{\partial x^2}.
\end{aligned}$$

We have

$$\begin{aligned}
\frac{\partial H_{m,n}}{\partial \eta} &= m\eta^{m-1}(1-\eta)^{n-m} - (n-m)\eta^m (1-\eta)^{n-m-1}, \\
&= \eta^{m-1}(1-\eta)^{n-m-1}\left[m(1-\eta) - (n-m)\eta\right],
\end{aligned}$$

and

$$\begin{aligned}
\frac{\partial^2 H_{m,n}}{\partial \eta^2} &= \frac{\partial}{\partial \eta}\left[m\eta^{m-1}(1-\eta)^{n-m} - (n-m)\eta^m(1-\eta)^{n-m-1}\right], \\
&= m\frac{\partial}{\partial \eta}\eta^{m-1}(1-\eta)^{n-m} - (n-m)\frac{\partial}{\partial \eta}\eta^m(1-\eta)^{n-m-1}, \\
&= m(m-1)\eta^{m-2}(1-\eta)^{n-m} - m(n-m)\eta^{m-1}(1-\eta)^{n-m-1} \\
&\quad -(n-m)m\eta^{m-1}(1-\eta)^{n-m-1} + (n-m)(n-m-1)\eta^m(1-\eta)^{n-m-2}.
\end{aligned}$$

Differentiating the average localization error with respect to the parameter x,

$$\begin{aligned}
\frac{\partial}{\partial x}\overline{\Omega(n)} &= \frac{\partial}{\partial x}\sum_{m=0}^{n}\binom{n}{m} H_{m,n}\Omega_{\boldsymbol{x},m} = \sum_{m=0}^{n}\binom{n}{m}\Omega_{\boldsymbol{x},m}\frac{\partial H_{m,n}}{\partial x}, \\
&= \left(\sum_{m=0}^{n}\binom{n}{m}\Omega_{\boldsymbol{x},m}\frac{\partial H_{m,n}}{\partial \eta}\right) \cdot \frac{\partial \eta}{\partial x}.
\end{aligned}$$

Therefore, When $\frac{\partial \eta}{\partial x} = 0$, $\frac{\partial}{\partial x}\overline{\Omega(n)} = 0$. Further,

$$\begin{aligned}
\frac{\partial^2 \overline{\Omega(n)}}{\partial x^2} &= \sum_{m=0}^{n}\binom{n}{m}\Omega_{\boldsymbol{x},m}\frac{\partial^2 H_{m,n}}{\partial x^2}, \\
&= \sum_{m=0}^{n}\binom{n}{m}\Omega_{\boldsymbol{x},m}\left(\left(\frac{\partial^2 H_{m,n}}{\partial \eta^2}\right) \cdot \left(\frac{\partial \eta}{\partial x}\right)^2 + \frac{\partial H_{m,n}}{\partial \eta} \cdot \frac{\partial^2 \eta}{\partial x^2}\right),
\end{aligned}$$

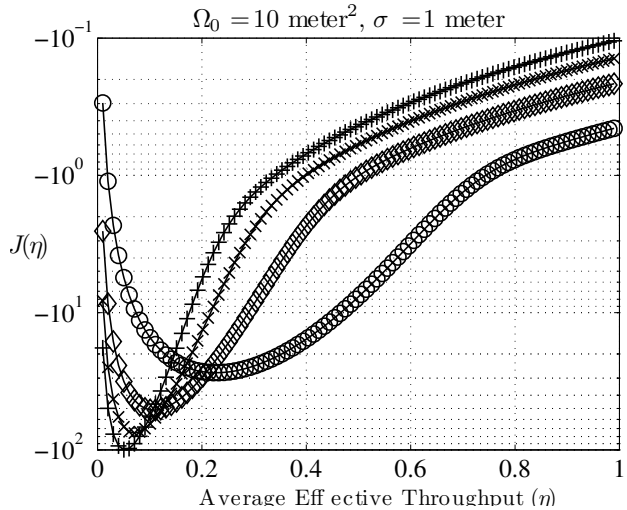


Figure 8.15: The variation of the $J(\eta)$ with average throughput η for different slot indices n. Here, $\Omega_0 = 10$ meter^2, and $\sigma = 1$ meter. We see that $J(\eta)$ is negative for all values of η.

When $\frac{\partial \eta}{\partial x} = 0$,

$$\frac{\partial^2 \overline{\Omega(n)}}{\partial x^2} = \left(\sum_{m=0}^{n} \binom{n}{m} \Omega_{\boldsymbol{x},m} \frac{\partial H_{m,n}}{\partial \eta} \right) \cdot \frac{\partial^2 \eta}{\partial x^2}.$$

When the throughput η is maximized with respect to a system parameter x, $\frac{\partial \eta}{\partial x} = 0$ and $\frac{\partial^2 \eta}{\partial x^2} < 0$, which implies:

$$\frac{\partial^2 \overline{\Omega(n)}}{\partial x^2} > 0, \text{ if } J(\eta) = \sum_{m=0}^{n} \binom{n}{m} \Omega_{\boldsymbol{x},m} \frac{\partial H_{m,n}}{\partial \eta} < 0,$$

$$\text{or, if } J(\eta) = \sum_{m=0}^{n} \binom{n}{m} \Omega_{\boldsymbol{x},m} \eta^{m-1} (1-\eta)^{n-m-1} \left[m(1-\eta) - (n-m)\eta \right] < 0.$$

Figure 8.15 shows the variation of $J(\eta)$ with η. We see that $J(\eta) < 0 \; \forall \eta$ when $\Omega_0 \gg \sigma^2$.

8.9 Appendix 8B: Statistics of the Interference Power from a Poisson Field of Interferers.

Suppose we define the "effective interference" as the random variable Y:

$$Y \triangleq \sum_k \frac{1}{r_k^{\beta}}. \tag{8.27}$$

Then Y represents the spatial dependence of the total interference seen from all interfering anchor nodes. Let $Y_{\epsilon a}$ be the interference power from anchors located in the annulus D_a between $r = \epsilon$ and $r = a$. The characteristic function of $Y_{\epsilon a}$ is given by

$$\phi_{Y_{\epsilon a}}(\omega) = E\left\{e^{j\omega Y_{\epsilon a}}\right\} = E_k\left\{E\left\{e^{j\omega Y_{\epsilon a}}|k \text{ interfering anchors in } D_a\right\}\right\}. \tag{8.28}$$

Given that there are k interfering anchors in D_a, due to the nature of the Poisson process [30], the distribution of their locations is that of k independent and identically distributed points with uniform distribution. Their distances from the origin are distributed as:

$$f_R(r) = \begin{cases} \frac{2r}{a^2-\epsilon^2} & \epsilon \le r \le a, \\ 0 & \text{otherwise.} \end{cases}$$

It follows that

$$E\left\{e^{j\omega Y_{\epsilon a}}|k \text{ interfering anchors in } D_a\right\} = \left(\int_\epsilon^a \frac{2r}{a^2-\epsilon^2} e^{\frac{j\omega}{r^\beta}} dr\right)^k.$$

From (8.4) and (8.28),

$$\begin{aligned}
\phi_{Y_{\epsilon a}}(\omega) &= \sum_{k=0}^{\infty} e^{-\Lambda'\pi(a^2-\epsilon^2)} \frac{(\Lambda'\pi(a^2-\epsilon^2))^k}{k!} \left(\int_\epsilon^a \frac{2r}{a^2-\epsilon^2} e^{\frac{j\omega}{r^\beta}} dr\right)^k \\
&= e^{-\Lambda'\pi(a^2-\epsilon^2)} \sum_{k=0}^{\infty} \frac{\left(\Lambda'\pi(a^2-\epsilon^2)\left(\int_\epsilon^a \frac{2r}{a^2-\epsilon^2} e^{\frac{j\omega}{r^\beta}} dr\right)\right)^k}{k!} \\
&= e^{-\Lambda'\pi(a^2-\epsilon^2)} e^{\left(\Lambda'\pi(a^2-\epsilon^2)\left(\int_\epsilon^a \frac{2r}{a^2-\epsilon^2} e^{\frac{j\omega}{r^\beta}} dr\right)\right)} = e^{\Lambda'\pi(a^2-\epsilon^2)\left(\int_\epsilon^a \frac{2r}{a^2-\epsilon^2} e^{\frac{j\omega}{r^\beta}} dr - 1\right)}.
\end{aligned}$$

From (8.29), it can be shown that [31] for $\beta > 2$, the characteristic function of Y as $a \to \infty$ and $\epsilon \to 0$ is given by

$$\begin{aligned}
\phi_Y(\omega) &= \lim_{\epsilon\to 0}\lim_{a\to\infty} \phi_{Y_a}(\omega) = \exp\left(jp\Lambda_L\pi\omega\left(\int_0^\infty t^{-\frac{2}{\beta}} e^{j\omega t} dt\right)\right) \\
&= \exp\left(-\pi\Lambda' e^{-\frac{\pi}{\beta}} \Gamma\left(1-\frac{2}{\beta}\right)\omega^{\frac{2}{\beta}}\right).
\end{aligned}$$

A closed-form expression for the probability density function (PDF) associated with the above characteristic function [85] is known only when $\beta = 4$. When $\beta = 2$, the interference power diverges. For $\beta > 2$, the PDF associated with the above characteristic function is given by:

$$f_Y(y) = \begin{cases} \frac{1}{\pi y}\sum_{k=1}^{\infty} \frac{\Gamma(\alpha k+1)}{k!}\left(\frac{\rho}{y^\alpha}\right)^k \sin\left(k\pi(1-\alpha)\right), & \beta > 2, \\ \frac{\pi\Lambda'}{2} y^{-\frac{3}{2}} e^{-\frac{\pi^3\Lambda'^2}{4y}}, & \beta = 4, \end{cases}$$

where $\alpha = \frac{2}{\beta}$, $\rho = \pi\Lambda'\Gamma(1-\alpha)$. The corresponding CDF is given by

$$F_Y(y) = \begin{cases} \frac{1}{\pi}\sum_{k=1}^{\infty}\frac{\Gamma(\alpha k)}{k!}\left(\frac{\rho}{y^{\alpha}}\right)^k \sin\left(k\pi(1-\alpha)\right), & \beta > 2 \\ \operatorname{erfc}\left(\frac{\pi^{\frac{3}{2}}\Lambda'}{2\sqrt{y}}\right), & \beta = 4. \end{cases} \tag{8.29}$$

8.10 Appendix 8C: Distribution of the SINR

Since the link distance to a localized node is also modeled as a random variable R, the random variable S representing the SINR seen at the unlocalized is given by

$$S = S(R,Y) = \left(\frac{1}{\xi_0} + \frac{K_1 R^{\beta} Y}{N_s}\right)^{-1}. \tag{8.30}$$

The random variables R and Y are assumed to be independent, and therefore, the expression for the CDF of the SINR conditioned on the link distance R is given by:

$$\begin{aligned} F_{S|R}(s|R) &= \Pr\{S \le s|R\} = 1 - \Pr\left\{Y \le \frac{1}{R^{\beta}K_1'}\left(\frac{1}{s} - \frac{1}{\xi_0}\right)\right\} \\ &= 1 - F_Y\left(\frac{1}{R^{\beta}K_1'}\left(\frac{1}{s} - \frac{1}{\xi_0}\right)\right). \end{aligned}$$

The probability of a packet being successfully decoded is modeled as the probability that the SINR exceeds the threshold ξ_T given a distance R, $P_s(R) = 1 - F_{S|R}(s = \xi_T|R)$, which implies

$$P_s(R) = F_Y\left(\frac{1}{R^{\beta}K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right). \tag{8.31}$$

Further,

$$f_{S|R}(s|R) = \frac{d}{ds}F_{S|R}(s|R) = \frac{f_Y\left(\frac{1}{R^{\beta}K_1'}\left(\frac{1}{s} - \frac{1}{\xi_0}\right)\right)}{R^{\beta}K_1' s^2}$$

Then,

$$\begin{aligned} f_{S|S\ge\xi_T,R}(s) &= \begin{cases} 0, & s < \xi_T, \\ \frac{f_{S|R}(s|R)}{1-F_{S|R}(s|R)(\xi_T)}, & s \ge \xi_T. \end{cases} \\ &= \begin{cases} 0, & s < \xi_T, s > \xi_0. \\ \frac{f_Y\left(\frac{1}{R^{\beta}K_1'}\left(\frac{1}{s}-\frac{1}{\xi_0}\right)\right)}{P_s(R)R^{\beta}K_1' s^2} & \xi_T \le s \le \xi_0. \end{cases} \end{aligned} \tag{8.32}$$

Average Probability of Success when $\beta = 4$

When $\beta = 4$, the expression for $P_s(R)$ becomes

$$P_s(R) = \text{erfc}\left(\frac{\pi^{\frac{3}{2}}\Lambda' R^2}{2\sqrt{\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}}\right).$$

Further, the probability of success averaged over the link distance is then given by

$$\overline{P_s} = \int_R P_s(r) f_R(r) dr. \tag{8.33}$$

When nodes are spatial Poisson distributed, the link distance can be modeled as a uniform random variable:

$$f_R(r) = \begin{cases} \frac{2r}{R_{\max}^2} & r \leq R_{\max}, \\ 0 & \text{otherwise.} \end{cases}$$

Then,

$$\overline{P_s} = \frac{2}{R_{\max}^2}\int_0^{R_{\max}} F_Y\left(\frac{1}{r^\beta K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right) r dr.$$

Let $y = \frac{r^{-\beta}}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)$. Therefore, $r^2 = y^{-\alpha}\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^\alpha$. This implies that $2rdr = -\alpha y^{-\alpha-1}\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^\alpha dy$. When $r = 0$, $y \to \infty$, and when $r = R_{\max}$, $y = \frac{R_{\max}^{-\beta}}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)$. Therefore,

$$\begin{aligned}
\overline{P_s} &= \frac{\alpha\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^\alpha}{R_{\max}^2}\int_{\frac{R_{\max}^{-\beta}}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}^{\infty} F_Y(y)\, y^{-\alpha-1} dy \\
&= \frac{\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^{1/2}}{2R_{\max}^2}\int_{\frac{R_{\max}^{-4}}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}^{\infty} \text{erfc}\left(\frac{\pi^{\frac{3}{2}}\Lambda'}{2\sqrt{y}}\right) y^{-3/2} dy \\
&= \frac{2\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^{1/2}}{\pi^{3/2}\Lambda' R_{\max}^2}\int_0^{\frac{\pi^{3/2}\Lambda' R_{\max}^2}{2\sqrt{\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}}} \text{erfc}(x)\, dx \\
&= \frac{\int_0^z \text{erfc}(x)\, dx}{z} = \text{erfc}\left(\frac{z}{2}\right) + 2\left(\frac{1 - e^{-\frac{z^2}{4}}}{\sqrt{\pi} z}\right),
\end{aligned}$$

where

$$z = \left(\frac{\pi\Lambda'\Gamma(1-\alpha)R_{\max}^2}{\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)^\alpha}\right). \tag{8.34}$$

Range Estimate Variances when $\beta = 4$

Assuming that a packet can be successfully decoded, the variance of range estimate that can be obtained

$$V = \frac{K_R}{S}. \tag{8.35}$$

assuming that the packets are received successfully, we have

$$\begin{aligned} F_{V|R}(v|R) &= \Pr\{V \le v|R\} = \Pr\left\{\frac{K_R}{S} \le v|S \ge \xi_T, R\right\} \\ &= \Pr\left\{S \ge \frac{K_R}{v}|S \ge \xi_T, R\right\} = 1 - F_{S|S\ge\xi_T,R}\left(\frac{K_R}{v}\right) \\ \Rightarrow \frac{d}{dv}F_{V|R}(v|R) &= f_{V|R}(v|R) = \frac{K_R}{v^2} f_{S|S\ge\xi_T,R}\left(\frac{K_R}{v}\right). \end{aligned}$$

This implies

$$f_{V|R}(v|R) = \frac{K_R}{v^2} \cdot \frac{f_Y\left(\frac{1}{R^\beta K_1'}\left(\frac{1}{\left(\frac{K_R}{v}\right)} - \frac{1}{\xi_0}\right)\right)}{P_s(R)R^\beta K_1'\left(\frac{K_R}{v}\right)^2} = \frac{f_Y\left(\frac{1}{R^\beta K_1'}\left(\frac{v}{K_R} - \frac{1}{\xi_0}\right)\right)}{P_s(R)R^\beta K_1' K_R}.$$

When $\beta = 4$,

$$f_{V|R}(v|R) = \frac{\frac{\pi\Lambda' R^2}{2}\left(\frac{1}{K_1'}\left(\frac{v}{K_R} - \frac{1}{\xi_0}\right)\right)^{-\frac{3}{2}} e^{-\frac{\pi^3 \Lambda'^2 R^4}{4\left(\frac{1}{K_1'}\left(\frac{v}{K_R} - \frac{1}{\xi_0}\right)\right)}}}{\operatorname{erfc}\left(\frac{\pi^{\frac{3}{2}}\Lambda' R^2}{2\sqrt{\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)}}\right) K_1' K_R}.$$

Let $a = \frac{\pi\Lambda' R^2}{2}$, $P_s(R) = \operatorname{erfc}\left(\frac{\pi^{\frac{3}{2}}\Lambda' R^2}{2\sqrt{\left(\frac{1}{K_1'}\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)\right)}}\right)$. Therefore,

$$\begin{aligned} \sigma^2(R) &= \int_{\frac{K_R}{\xi_0}}^{\frac{K_R}{\xi_T}} v f_{V|R}(v|R) dv \\ &= \frac{a}{P_s(R)K_1' K_R} \int_{\frac{K_R}{\xi_0}}^{\frac{K_R}{\xi_T}} v \cdot \left(\frac{1}{K_1'}\left(\frac{v}{K_R} - \frac{1}{\xi_0}\right)\right)^{-\frac{3}{2}} e^{-\frac{\pi a^2}{\frac{1}{K_1'}\left(\frac{v}{K_R} - \frac{1}{\xi_0}\right)}} dv. \end{aligned}$$

Let $y = \frac{\pi a^2 K_1'}{\left(\frac{v}{K_R} - \frac{1}{\xi_0}\right)}$. This implies $v = K_R\left(\frac{\pi a^2 K_1'}{y} + \frac{1}{\xi_0}\right)$. This implies: $dv = -\frac{K_R \pi a^2 K_1'}{y^2} dy$.

$$\begin{aligned}
\sigma^2(R) &= \frac{K_R}{\sqrt{\pi} P_s(R)} \int_{\frac{\pi a^2 K_1'}{\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}}^{\infty} \left(\frac{\pi a^2 K_1'}{y} + \frac{1}{\xi_0}\right) \cdot y^{-\frac{1}{2}} e^{-y} dy \\
&= \frac{K_R}{\sqrt{\pi} P_s(R)} \left[\pi a^2 K_1' \Gamma\left(-\frac{1}{2}, \frac{\pi a^2 K_1'}{\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}\right) + \frac{1}{\xi_0} \Gamma\left(\frac{1}{2}, \frac{\pi a^2 K_1'}{\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}\right)\right].
\end{aligned}$$

If $x = \frac{\pi a^2 K_1'}{\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)} = \frac{\pi^3 \Lambda'^2 R^4 K_1'}{\left(\frac{1}{\xi_T} - \frac{1}{\xi_0}\right)}$,

$$\Gamma\left(-\frac{1}{2}, x\right) = 2\left[\frac{\exp(-x)}{\sqrt{x}} - \sqrt{\pi}\mathrm{erfc}(\sqrt{x})\right], \tag{8.36}$$

and

$$\Gamma\left(\frac{1}{2}, x\right) = \sqrt{\pi}\mathrm{erfc}(\sqrt{x}), \tag{8.37}$$

which leads to

$$\sigma^2(R) = \frac{K_R}{\sqrt{\pi} P_s(R)} \left[\pi a^2 K_1' \cdot 2\left(\frac{\exp(-x)}{\sqrt{x}} - \sqrt{\pi}\mathrm{erfc}(\sqrt{x})\right) + \frac{1}{\xi_0}\sqrt{\pi}\mathrm{erfc}(\sqrt{x})\right].$$

The range estimate variance averaged over the link distance can be obtained using the following expression via numerical integration:

$$\overline{\sigma^2} = \int_R \sigma^2(R) f_R(r) dr.$$

Chapter 9

Power Control for UWB PoLoNets

9.1 Introduction

In this chapter, we consider the problem of designing power-control algorithms for UWB PoLoNets. As a mobile node moves through a network of localized reference nodes, its localization accuracy fluctuates due to variations in connectivity, range estimate variances and geometry. Further, if the probability of reference node failure is significant, controlling the transmit power to ensure the minimum required connectivity becomes essential. We demonstrate that power-control algorithms can be used to improve and provide robustness to a mobile node's localization accuracy as it traverses the network of reference nodes, without the necessity to transmit at the maximum available power. Additionally, as suggested in Chapter 5, efficient power-control can speed up the rate at which range estimates are delivered to unlocalized nodes, thereby promoting rapid localization.

Key contributions included in this chapter are:

- While power-control algorithms have been developed for cellular networks, their design is typically from the perspective of ensuring requisite signal-to-interference-and-noise ratios or bit/frame error-rates. Power-control from the perspective of localization accuracy has not been considered previously.
- An analytical framework for the impact of power-control on localization accuracy is provided.
- Two different power-control algorithms based on localization accuracy have been proposed, whose efficacy is demonstrated via simulation results.

This chapter is organized as follows: In sections 9.2 and 9.3, we examine the variation of localization accuracy due to node mobility and impact of power control on the localization

accuracy. A means of quantifying and estimating the "quality" of location-estimates is discussed in section 9.4. Two power control estimates based on the defined quality metric are discussed in detail in section 9.5. A comparison of these power control approaches with simulated optimal solutions obtained can be found in section 9.6. We conclude in section 9.7.

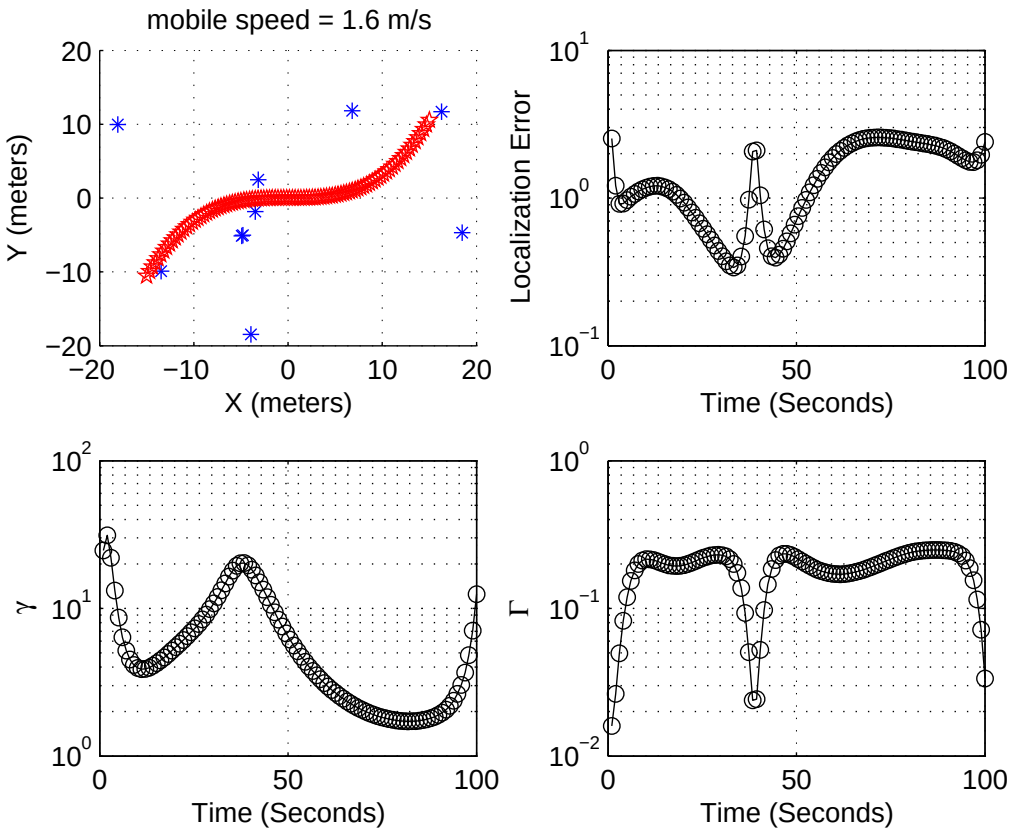


Figure 9.1: "Spatial Fading" of Localization Accuracy: the mobile is assumed to be moving through the $L \times L$ area of interest with speed $v = 1.6$ meters/second. The N localized reference nodes are assumed to uniformly (randomly) distributed over the area of interest. In the above example, $\beta = 2$, $P = 10$ mW, $N = 10$, $L = 40$ meters, $\xi_{\min} = 20$ dB.

9.2 "Fading" of Localization Accuracy

As shown in (3.5), the localization error $\Omega_{\boldsymbol{x}}$ is dependent on the geometry of reference nodes (through Γ_m), connectivity with localized reference nodes (through m) and range estimate variances (which depend on R_i). It is evident that as a mobile node moves through

the network of reference nodes, all three quantities can vary with time and therefore, the localization error of a mobile node is a function of time. Figure 9.1 illustrates an example of the variation of a mobile node's localization error as the mobile node moves through an area containing randomly distributed reference nodes. The localization error fluctuation is analogous to the spatial fading of received signal power in wireless propagation channels.

It was shown in (3.11) that increasing the number of range estimates from reference nodes decreases the localization error. Therefore, one way of reducing the fluctuation of localization error due to changes in connectivity is to initially deploy reference nodes with a high node density. An alternative means of improving localization accuracy and making it robust to node mobility is through the use of power control. We see that increasing the transmit power P (i) increases the transmission range $R_{\max}$, improving connectivity and resulting in the availability of a larger number of range estimates, and (ii) from (2.25), reduces the range estimate variances. In the following section, we quantify the decrease in localization error with an increase in transmit power.

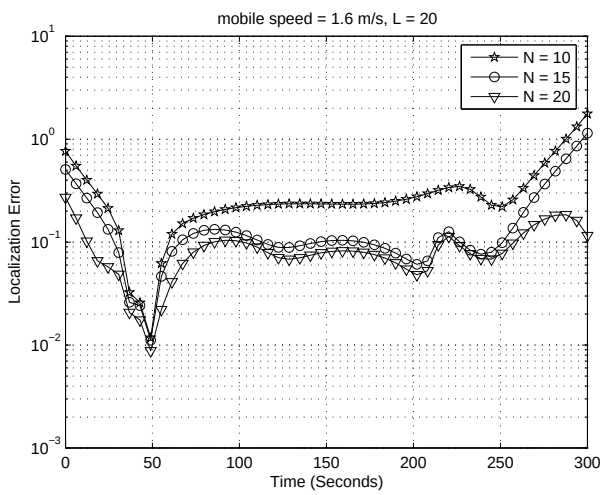


Figure 9.2: Localization Accuracy "Fading" versus node density: the number of localized reference nodes N in the $L \times L$ area of interest is increased. We see that as the average localized node density $\rho = \frac{N}{L^2}$ increases, a mobile's localization accuracy is more robust to fluctuations. In the above simulation, $L = 20$ meters, and the mobile speed is $v = 1.6$ meters/second. For a given set of anchor locations, the number of available anchors is increased from $N = 10$ to $N = 20$.

It must be pointed out that the short-term fading in localization accuracy can be partially mitigated by tracking algorithms such as Kalman filtering as will be seen in Chapter 6. However, in PoLoNets with a low density of reference nodes, limited connectivity with reference nodes — which results in poor localization accuracy — can span longer durations of time, thereby limiting the effectiveness of tracking algorithms. In such situations, we can

resort to power-control to improve connectivity, reduce battery consumption, reduces interference, and reduces localization error. In the following section, we quantify the decrease in localization error with an increase in transmit power. It must be noted that although the investigation of power control algorithms that follows is performed with UWB PoLoNets in mind, the development is not specific to the UWB physical layer and be applied to PoLoNets with a generic physical layer.

9.3 Effect of Increasing Transmit Power

Let P be the transmit power of a mobile node. The effective received signal-to-noise-ratio (SNR) at reference node i, during a "range-initiate" broadcast by the mobile node is modeled by $\xi_i = K_P P R_i^{-\beta}$, where β is the path-loss exponent, $R_i = \|\boldsymbol{x} - \boldsymbol{x}_i\|$ is the distance between the nodes, and K_P is a constant that subsumes the effects of other physical layer parameters. If multiple uncoordinated mobile nodes are present, then the SNR would have to be replaced by signal-to-interference-and-noise ratio (SINR). For simplicity, we temporarily examine the case with one mobile node in a network of reference nodes. We assume that the minimum SNR required for a packet to be received successfully is given by $\xi_{\min}$, resulting in a maximum radius of coverage $R_{\max}(P) = \left(K_P P \xi_{\min}^{-1}\right)^{\frac{1}{\beta}}$ around $\boldsymbol{x}$. Let $N(P)$ represent the number of reference nodes present within a radius $R_{\max}(P)$ from the mobile node.

Given $m = N(P)$ range estimates from reference nodes, the mobile node estimates its location with the localization error lower-bounded by (3.5). Suppose the transmit power is increased by a factor $\chi > 0$. The new transmit power is given by $P' = P\,(1+\chi)$. In the absence of multiple-access interference, the number of reference nodes that can receive a packet from the mobile node increases (or remains the same) with this increase in the transmit power. Let the new number of responding reference nodes be $N(P') = m + n \geq m$. From (2.25), the new range estimate variances are given by

$$\sigma_i'^2 = \frac{K_R R_i^{\beta}}{K_P P'} = \frac{\sigma_i^2}{1+\chi}, \qquad i = 1, 2, \cdots, m+n. \tag{9.1}$$

The improvement in terms of localization error is defined as $D = \Omega_{\boldsymbol{x}}(P) - \Omega_{\boldsymbol{x}}(P')$. The expression for D in terms of χ is derived in Appendix 9A and is given by (9.10). It is clear from (9.10) that $D > 0$ if $P' > P$ which implies that the localization error decreases with an increase in transmit power. Figure 9.3 illustrates the impact of increasing the transmit power on localization accuracy obtained via (a) the CRLB, and (b) the LS estimator. We see that under ideal conditions, increasing the transmit power can result in significant improvement in localization accuracy. As stated previously, the decrease in the localization error can be attributed to the two effects of increasing the transmit power: improved connectivity, resulting a larger number of range estimates ($m + n \geq m$), and reduced range variances ($\sigma_i'^2 < \sigma_i^2$).

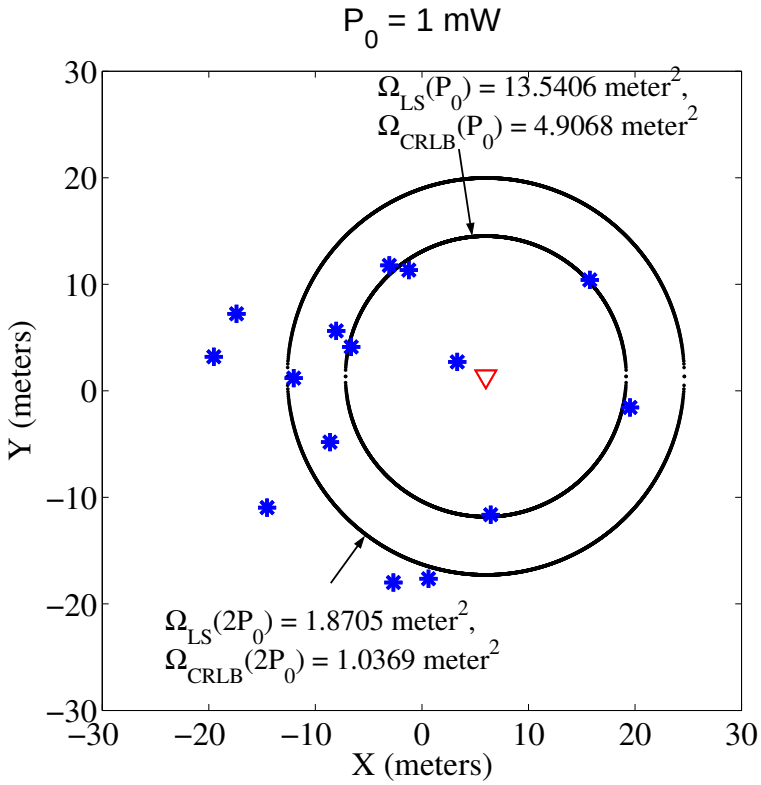


Figure 9.3: Impact of Power Control: Increasing the transmit power from $P_0 = 1$ mW to $2P_0$ (a) increases the number of available range estimates from $m = 4$ to $m = 9$, and (b) reduces the variance of range estimates. This leads to a decrease in the average localization error. In the above example, $\boldsymbol{x} = [5\ 5]^T$, and $N_R = 15$ reference nodes are randomly distributed over an $L \times L$ area of interest, where $L = 40$ meters.

9.4 Quality of Location Estimates

We define the *Quality of Localization* (QOL) $Q_{\boldsymbol{x}}$ of a location-estimate $\boldsymbol{x}$ as the *reciprocal of the localization error*:

$$Q_{\boldsymbol{x}}(P) = \frac{1}{\Omega_{\boldsymbol{x}}} = \frac{K_P P \sum_{i=1}^{m} \sum_{j=1, j>i}^{m} \frac{\sin^2(\alpha_i - \alpha_j)}{R_i^\beta R_j^\beta}}{K_R \sum_{i=1}^{m} \frac{1}{R_i^\beta}}. \tag{9.2}$$

The larger the value of the QOL, the higher the localization accuracy. It is evident that in an infinite field of reference nodes as $P \to \infty$, $Q_{\boldsymbol{x}}(P) \to \infty$. We would like to find the transmit power P with which a given target location-estimate quality Q_0 is obtained. We can therefore set up an objective function that is to be minimized as

$$Z(P) = \left(Q_{\boldsymbol{x}}(P) - Q_0\right)^2. \tag{9.3}$$

It can be verified from (9.2) that the transmit power $P = P^\star$ that minimizes this objective function is given by the solution of the following equation:

$$P^\star = \frac{Q_0 K_R \sum_{i=1}^{N(P^\star)} \frac{1}{R_i^\beta}}{K_P \sum_{i=1}^{N(P^\star)} \sum_{j=1, j>i}^{N(P^\star)} \frac{\sin^2(\alpha_i - \alpha_j)}{R_i^\beta R_j^\beta}}. \tag{9.4}$$

However, this equation cannot directly be used to solve for $P^\star$, since we do not *a priori* know the values of $N(P)$, $\{R_i\}$ or $\{\alpha_i\}$ for a specific mobile location in the network. In our case, the value of $P^\star$ can be obtained via simulations in order to serve as a benchmark for the power control schemes proposed in the following sections.

9.4.1 Quality Metric

In order to implement power control algorithm based on the quality of location estimates, it is essential for a mobile node to be able to assess the quality of its location-estimate. Suppose an initial power level P_0 was used and a location estimate $\hat{\boldsymbol{x}}$ was obtained (e.g., using a least-squares estimator). Assuming the knowledge of the constants K_P and K_R via calibration, the range estimates r_i, and the reference node coordinates $\boldsymbol{x}_i$, $i = 1, 2, \cdots, m$, the quality of the location estimate can be estimated as follows: the angles $\hat{\alpha}_i$ can be estimated using $\hat{\alpha}_i = \measuredangle(\boldsymbol{x}_i - \hat{\boldsymbol{x}})$. The variances of the range estimates can be estimated using

$$\hat{\sigma}_i^2 = \frac{K_R r_i^\beta}{K_P P_0}. \tag{9.5}$$

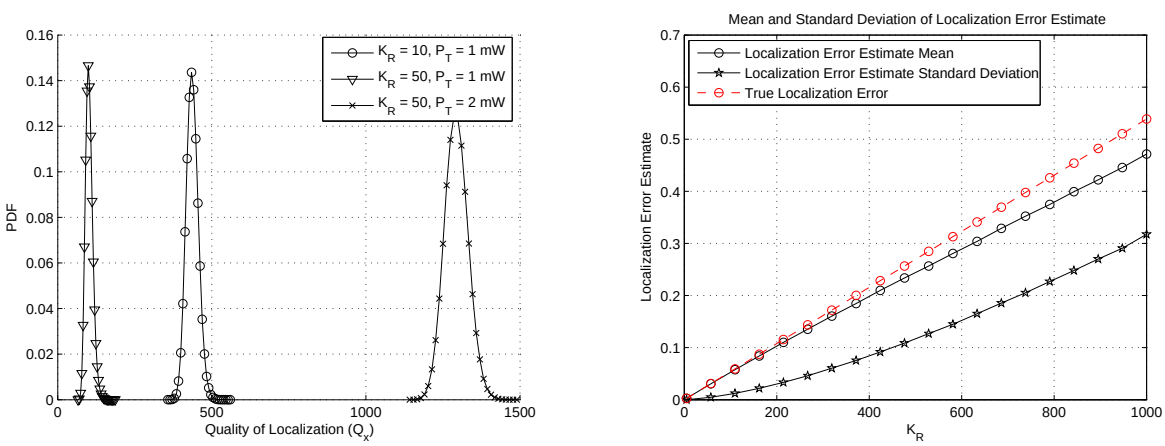
Therefore, an estimate $\hat{Q}_{\hat{\boldsymbol{x}}}$ of the QOL can be computed using

$$\hat{Q}_{\hat{\boldsymbol{x}}} = \frac{\sum_{i=1}^{m} \sum_{j=1, j>i}^{m} \frac{\sin^2(\hat{\alpha}_i - \hat{\alpha}_j)}{\hat{\sigma}_i^2 \hat{\sigma}_j^2}}{\sum_{i=1}^{m} \frac{1}{\hat{\sigma}_i^2}}. \tag{9.6}$$

A mobile node can then adapt its transmit-power based on the QOL estimates as described in the following sections. Figure 9.4(a) shows the distribution of $\hat{Q}_{\hat{\boldsymbol{x}}}$ computed from Gaussian range estimates and a least-squares (LS) estimate of the nodes location. The estimate $\hat{Q}_{\hat{\boldsymbol{x}}}$ can be approximated by a Gaussian random variable, especially for large K_R. We further observe from Figure 9.4(a) that the mean and standard deviation of the QOL estimate increase with K_R. Figure 9.4(b) shows that the QOL estimate described in (9.6) can be treated as an unbiased estimate of the true QOL for small values of K_R:

$$\hat{Q}_{\hat{\boldsymbol{x}}} = Q_{\boldsymbol{x}} + v_Q, \tag{9.7}$$

where $v_Q \sim \mathcal{N}(0, \sigma_Q^2(P))$, where $\sigma_Q^2(P)$ is a non-increasing function of P. In the following section, we define an iterative power-control algorithm based on the QOL estimate defined in (9.6).



(a) (b)

Figure 9.4: (a) The histograms of $\hat{Q}_{\hat{\boldsymbol{x}}}$ for different values of K_R, and (b) the mean and standard deviation of $1/\hat{Q}_{\hat{\boldsymbol{x}}}$ for different values of K_R for a given geometry of reference nodes and a large number of range measurement noise realizations. The true value of $1/Q_{\boldsymbol{x}}$ is shown for the sake of comparison.

9.5 Transmit power adaptation based on quality of location estimates

In this section, we describe two power control algorithms for UWB PoLoNets based on location-estimates in order to obtain a target QOL Q_0. The first algorithm uses an adaptive

power-update based on the estimated QOL, whereas the second is a two-step procedure that tries to directly compute the transmit power required to obtain the target localization accuracy. We assume that mobile node displacements are small in the time-span over which power control algorithms are implemented.

9.5.1 QOL-based Iteration

In a mobile environment, it may be necessary to update the location-estimate of the mobile node periodically. In such a case, the transmit power can be updated at each attempt to compute the location-estimate, based on the quality of the previous location estimate. Consider a simple iterative power control algorithm, where the kth iteration is of the form

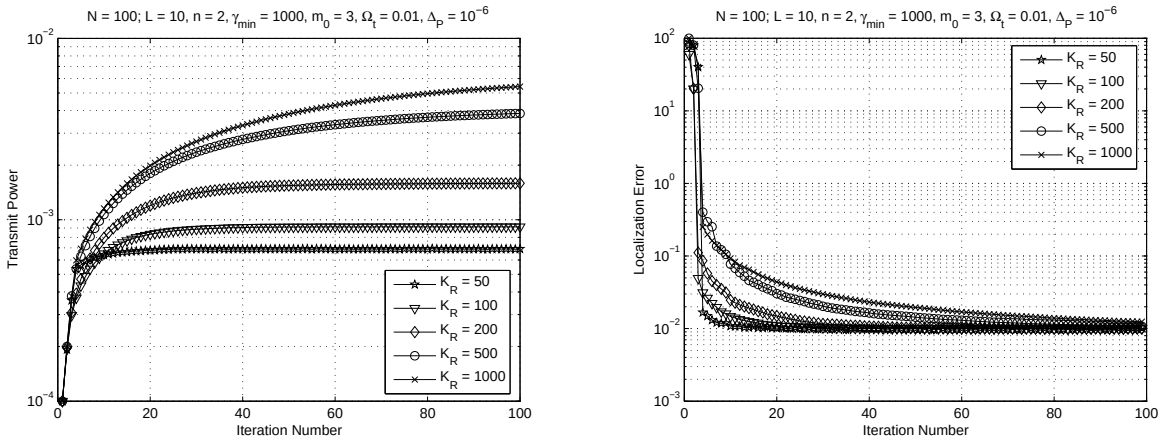
$$P_{k+1} = P_k + \Delta_{P,k} \left(Q_0 - \hat{Q}_{\hat{\boldsymbol{x}}}(P_k) \right), \tag{9.8}$$

where $\Delta_{P,k}$ represents the step-size of the power increment or decrement in the kth iteration. It is clear from the above equation that if $\hat{Q}_{\hat{\boldsymbol{x}}}(P_k) < Q_0$, then the transmit power in the subsequent iteration is increased; if $\hat{Q}_{\hat{\boldsymbol{x}}}(P_k) > Q_0$, the transmit power is decreased. Based on the framework of power control convergence analysis provided in [32], [33], it can be shown that a power control scheme of a similar form as (9.8) converges with probability one to the optimal transmit power, provided the value of $\Delta_{P,k}$ and K_R are sufficiently small. The proof is provided in Appendix 9B, where we assume that $\hat{Q}_{\hat{\boldsymbol{x}}}$ is an unbiased estimate of the QOL, an assumption that is valid for small values of K_R.

Figures 9.5-9.6 show the convergence of the transmit power and localization error using the above scheme for different values of K_R. It is clear that as K_R increases (resulting in the increase of range variances), a larger step-size $\Delta_{P,k} = \Delta_P$ results in faster convergence of the power-control algorithm. If K_R is known *a priori*, then the speed of convergence of the above iterative scheme to the desired QOL Q_0 can be increased by making Δ_P proportional to K_R as seen in Figure 9.6.

9.5.2 Two-Step Solution

In applications where a location-estimate is computed at irregular intervals or based on requirement, an iterative scheme such as that described above may not be suitable. In such a case, we would like a non-iterative solution to directly compute the required transmit power. Let P_0 be the initial transmit power with the corresponding QOL estimate $\hat{Q}_{\hat{\boldsymbol{x}}}(P_0)$. From (9.10), the unknown increment χ can be computed (as shown in Appendix 9A) setting $D = \frac{1}{\hat{Q}_{\hat{\boldsymbol{x}}}(P_0)} - \frac{1}{Q_0}$, if the term C in (9.10) is known. However, C includes range estimates from nodes that *would* be responding *after* the transmit power is increased by a factor χ. Since this information is not known when χ is computed, we assume a worst-case geometry



(a) (b)

Figure 9.5: (a) Transmit power, and (b) localization error versus iteration index for the QOL-based iterative scheme in (9.8). In this case a constant step-size Δ_P is used, which, along with other parameters is displayed above the figure. In the above figures, the desired QOL is $Q_0 = 100$ meters^{-2}.

for nodes that would be introduced if the transmit power were increased, given by (9.11), leading to $C = 0$. The details of the computation of χ can be found in Appendix 9A.

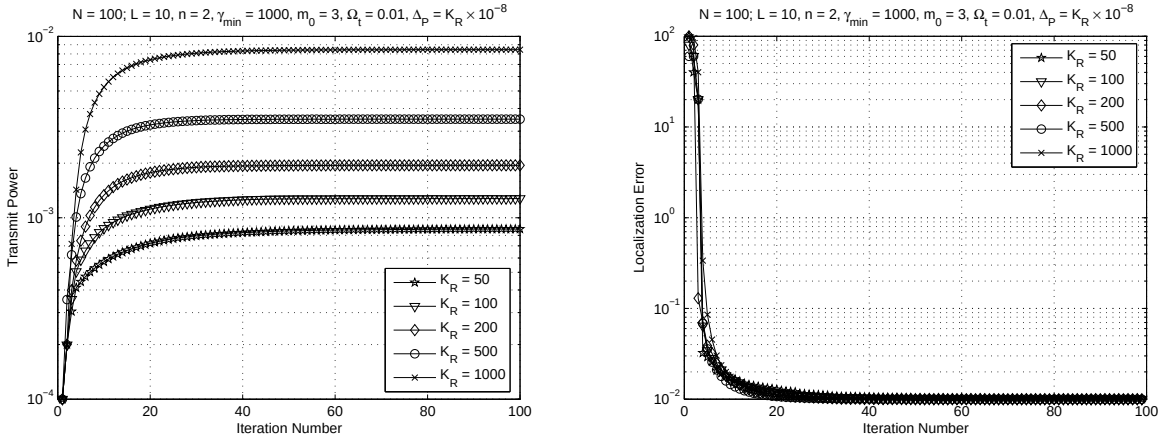
Since the computed value of χ assumes a worst-case geometry, the transmit power $P' = P_0(1+\chi)$ is, with high probability, larger than the optimal transmit power $P^\star$ that ensures a quality metric Q_0. Therefore, the second step in the solution is to decrease the transmit power to a level that maintains $\hat{Q}_{\hat{\boldsymbol{x}}} \approx Q_0$. This ensures that if the location-estimate is updated in quick succession subsequently, the initial transmit power used is close to the optimal transmit power. In this case, since the locations and ranges for all the nodes are known (from the first step), we can decrease the transmit power so as to selectively exclude the largest range estimates and retain only the range estimates that ensure $\hat{Q}_{\hat{\boldsymbol{x}}} \approx Q_0$.

The above power control algorithm is defined explicitly as follows:

(1) Start with a transmit power P_0.

(2) Using transmit power P_0, obtain range estimates r_i, $i = 1, 2, \cdots, m$ and reference node coordinates $\boldsymbol{x}_i$, $i = 1, 2, \cdots, m$. Compute an estimate of the mobile node's location $\hat{\boldsymbol{x}}$. Using this location estimate, estimate the angles $\hat{\alpha}_i = \measuredangle(\boldsymbol{x}_i - \hat{\boldsymbol{x}})$. Compute the range variances using (9.5). Compute the QOL the location-estimate $\hat{\boldsymbol{x}}$ using (9.6).

(3) If $\hat{Q}_{\hat{\boldsymbol{x}}} > Q_0$, then arrange the available range estimates in increasing order. Compute the QOL estimate from (9.6) with the first $m = k$ ranges, ending with the full set of ranges.



(a) (b)

Figure 9.6: (a) Transmit power, and (b) localization error versus iteration index for the QOL-based iterative scheme in (9.8). In this case, the step-size Δ_P is varied according to K_R which determines the variance of the range estimates. In the above figures, the desired QOL is $Q_0 = 100$ meters^{-2}.

Find the value of $k = k^\star$ for which the quality metric $\hat{Q}_{\hat{\boldsymbol{x}}}(k)$ is closest to Q_0. Compute the transmit power corresponding to the $k^\star$th range estimate using $P = \frac{r_{k^\star}^{\beta}\xi_{\min}}{K_P}$. Stop.

(4) Else, if $\hat{Q}_{\hat{\boldsymbol{x}}} < Q_0$, compute χ using (9.12), where we replace $\{\sigma_i^2\}$ with $\{\hat{\sigma}_i^2\}$, $\{\alpha_i\}$ with $\{\hat{\alpha}_i\}$, and D with $\left(\frac{1}{Q_0} - \frac{1}{\hat{Q}_{\hat{\boldsymbol{x}}}}\right)$. Next, go to step (3).

It must be noted that in general, the transmit power obtained using step (3) will not be equal to the optimal transmit power because (i) there may not exist a $k^\star$ such that $\hat{Q}_{\hat{\boldsymbol{x}}}(k^\star) = Q_0$, (in such a case, the final $k^\star = \text{argmin}_k|\hat{Q}_{\hat{\boldsymbol{x}}}(k) - Q_0|$), and (ii) the transmit-power is computed using $P = \frac{r_{k^\star}^{\beta}\xi_{\min}}{K_P}$, which uses a range estimate that is a noisy estimate of the distance between the nodes.

9.6 Simulation Results

Figure 9.7 compares the performance of the two power control schemes described in the previous section with the optimal solution $P^\star$ obtained through (9.4) for a given mobile location. For different values of N_R and K_R (which controls the variance of the range estimates through (2.25)), we observe that the transmit power obtained by both schemes closely matches the optimal transmit power $P^\star$ in all cases. The value of L was fixed at 10

meters, and the number of randomly distributed reference nodes N_R was increased, thereby increasing the effective node density. It is observed that the transmit power required to attain a QOL $Q_0 = 100\ m^{-2} \Rightarrow \Omega_0 = 0.01\ m^2$ decreases as the node density increases, which is expected from (3.11), since for a given transmit power, the number of available range estimates m increases as we increase the node density. Also, as K_R increases, the range estimate variances increase and we require a higher transmit power to attain the target localization error $\Omega_0 = 1/Q_0$.

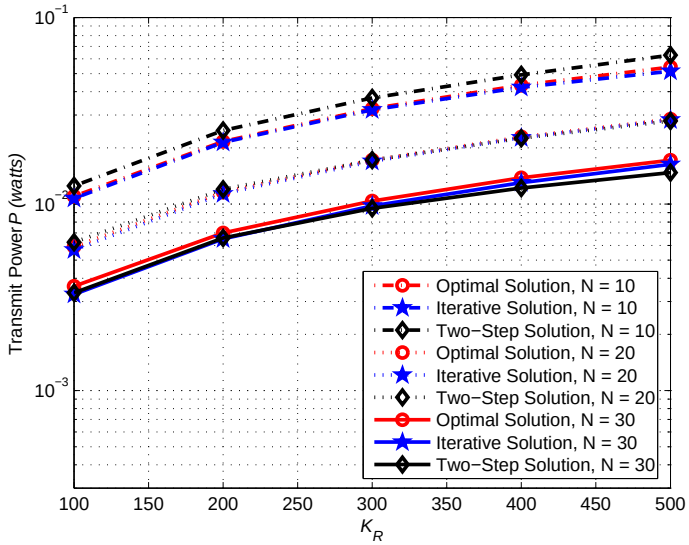


Figure 9.7: Comparison of the two power control schemes with the optimal solution: in this simulation, $L = 10$, $\beta = 2$, $\xi_{min} = 20$ dB, $Q_0 = 100\ m^{-2}$.

Figure 9.8 shows the impact of using iterative power control on a mobile node's localization accuracy as it moves through a network of $N_R = 15$ reference nodes that are randomly scattered over a 40×40 meter^2 area. The mobile node is assumed to move at a speed of $v_m = 1.6\ m/s$ and $N_{iter} = 10$ power control iterations are performed per second. We observe that the use of power control not only reduces the fluctuation of localization accuracy, but also ensures lower localization error, i.e., higher localization accuracy. However, this gain comes at the price of using a higher transmit power, which would lead to higher interference levels when several mobile nodes are present, which impacts ranging accuracy.

9.6.1 Power-control with multiple mobile nodes

In the previous sections, we looked at power-control schemes considering the localization accuracy of a single unlocalized node. When there are K unlocalized nodes simultaneously ranging to N localized nodes, increasing the transmit power of each unlocalized nodes can

adversely impact the localization accuracy of other unlocalized nodes due to the increase in multiple-access interference. In such a case, the power control algorithm needs to be suitably designed to provide the desired QOL for all unlocalized nodes. Suppose $\boldsymbol{p}$ is the $K \times 1$ vector of transmit powers of unlocalized nodes. Then, the SINR at the lth unlocalized node receiving a range packet from the jth localized node with transmit power P_l can be written as a function of the transmit powers:

$$\xi_{l,j}(\boldsymbol{p}) = \frac{P_l R_{lj}^{-n}}{\frac{1}{N_s}\sum_{k=1,k\neq l}^{K} P_k R_{lk}^{-n} + P_N}, \; l \in \{1, 2, \cdots, K\}, \; j \in \{1, 2, \cdots, N\},$$

where N_s accounts from the processing gain of the system, and P_N denotes the effective thermal noise power. Assuming that the interference from other mobile users is Gaussian distributed, the variance of the range estimate at mobile l from reference node j is given by

$$\sigma_{l,j}^2(\boldsymbol{p}) = \frac{K_R}{\xi_{i,j}} = \frac{K_R R_{lj}^{n}\left(\frac{1}{N_s}\sum_{k=1,k\neq l}^{K} P_k R_{lk}^{-n} + P_N\right)}{P_l}.$$

As the range estimate variances $\sigma_{l,j}^2$ and the number of range estimates available to the lth mobile node m_l are functions of the transmit power vector $\boldsymbol{p}$, we can express the quality of the location estimate of the lth mobile node as:

$$\begin{aligned} Q_l(\boldsymbol{p}) &= \frac{\sum_{j=1}^{m_l(\boldsymbol{p})}\sum_{i=1}^{m_l(\boldsymbol{p})} \frac{\sin^2(\alpha_i-\alpha_j)}{\sigma_{l,j}^2(\boldsymbol{p})\sigma_{l,i}^2(\boldsymbol{p})}}{\sum_{j=1}^{m_l(\boldsymbol{p})} \frac{1}{\sigma_{l,j}^2(\boldsymbol{p})}} \\ &= \frac{P_l}{K_R\left(\frac{1}{N_s}\sum_{k=1,k\neq l}^{K} P_k R_{lk}^{-n} + P_N\right)} \left(\frac{\sum_{j=1}^{m_l(\boldsymbol{p})}\sum_{i=1}^{m_l(\boldsymbol{p})} \frac{\sin^2(\alpha_i-\alpha_j)}{R_{lj}^{-n}R_{li}^{-n}}}{\sum_{j=1}^{m_l(\boldsymbol{p})} R_{lj}^{-n}}\right). \end{aligned}$$

As with the single-mobile case, we can define a power vector update:

$$\boldsymbol{p}_{k+1} = \boldsymbol{p}_k + \boldsymbol{\Delta}_{P,k}\left(Q_0 - \hat{\mathbf{Q}}(\boldsymbol{p}_k)\right), \tag{9.9}$$

where $\boldsymbol{\Delta}_{P,k}$ and $\hat{\mathbf{Q}}(\boldsymbol{p}_k)$ are respectively, the vector of step-sizes and quality estimates during iteration k. The main issue with such a *distributed* power control scheme is that several mobile users may be driven to increase their transmit powers without necessarily improving their respective localization accuracies while degrading the localization accuracies of other nodes. Further, proving stochastic convergence for the above iteration along the lines of the single-mobile scenario (Appendix 9B) is a difficult task. However, Figure 9.9 presents simulation results implementing the iterative power control scheme in (9.9) that indicate that similar schemes applied to multiple unlocalized nodes could result in higher localization accuracies. In Figure 9.9 the quality estimates were computed according to (9.6) and the iteration in (9.9) was applied with a constant step-size Δ_P and transmit power limits of $\{0.1, 5\}$ mW. Location-estimates were obtained using the LS estimator. We see that the application of such a power-control method can provide improved average localization accuracy over time.

9.7 Conclusions

In this chapter, we looked at the use of power control schemes in UWB PoLoNets. The localization accuracy of a mobile node was shown to be determined by connectivity with reference nodes, range estimate variances and geometry of mobile nodes relative to the reference nodes. This results in large-scale fluctuation of localization accuracy as the mobile node traverses the network. Two ameliorative power control schemes for a single unlocalized node were discussed: a stochastically convergent iterative scheme based on the quality of location-estimates and a two-step analytical solution. We demonstrated that the two approaches approximate the empirically-obtained optimal solution, and significantly improve the robustness of location-estimates. Although the stochastic convergence of the iterative power control algorithm was not analytically shown in the case of multiple unlocalized nodes, its efficacy in such scenarios was demonstrated via simulations.

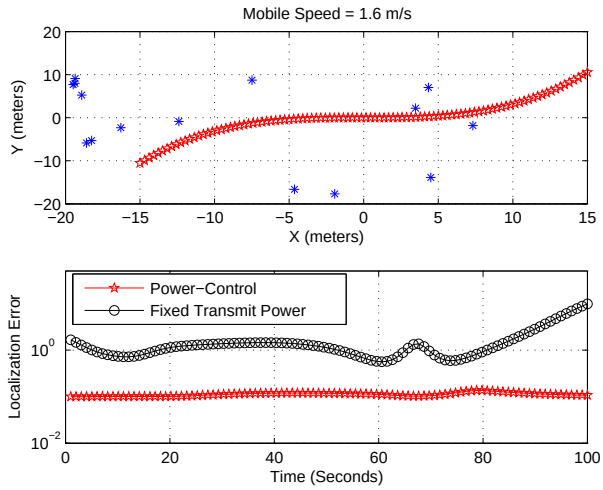


Figure 9.8: (a) Trajectory of a mobile node, and (b) reduction in "spatial fading" of localization error using power control. The parameters used for this simulation were: $P_0 = 5$mW, $\beta = 2$, $N = 25$, $L = 40$, $\xi_{\min} = 20$ dB, $v_m = 1.6$ m/s.

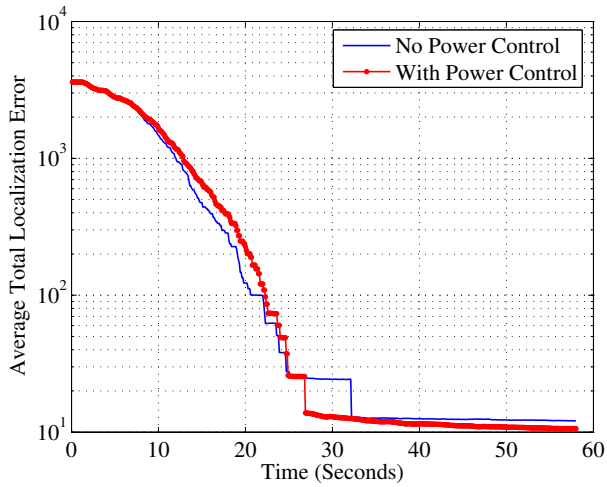


Figure 9.9: Simulation of an iterative power-control algorithm with multiple unlocalized nodes. Here, $K = 20$, $N = 5$, $L = 20$ meters, $\beta = 2$, $Q_0 = 0.1$ meter^{-2}. The unlocalized nodes are assumed to be stationary for simplicity. The time-varying localization error is averaged over a large number of realizations of unlocalized node locations.

9.8 Appendix 9A: Variation of Ω_x with Transmit Power

Let us define the following terms:

$$\begin{aligned} g_m &= \sqrt{1-4\Gamma_m},\ \gamma_n = \sum_{i=m+1}^{m+n} \frac{1}{\sigma_i^2} = \gamma_{m+n} - \gamma_m, \\ \psi_n &= \sum_{i=m+1}^{m+n} \sum_{j=m+1, j>i}^{m+n} \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2 \sigma_j^2},\ \Gamma_n = \frac{\psi_n}{\gamma_n^2}. \end{aligned}$$

We have

$$\Omega(P') = \frac{\sum_{i=1}^{m} \frac{1}{\sigma_i'^2}}{\sum_{i=1}^{m} \sum_{j=1, j>i}^{m} \frac{\sin^2(\alpha_i - \alpha_j^2)}{\sigma_i'^2 \sigma_j'^2}} = \frac{\gamma_{m+n}}{(1+\chi)\,\psi_{m+n}}.$$

The denominator term is simplified using

$$\begin{aligned} \psi_{m+n} &= \sum_{i=1}^{m+n} \sum_{j=1, j>i}^{m+n} \frac{\sin^2(\alpha_i - \alpha_j)}{\sigma_i^2 \sigma_j^2} \\ &= \psi_m + \frac{\gamma_m}{2} \sum_{i=m+1}^{n} \frac{1}{\sigma_i^2} \left[1 - g_m \cos(2\alpha_i - 2\nu)\right] + \psi_n \\ &= \psi_m + \frac{\gamma_m}{2} \left[\gamma_n - g_m \sum_{i=m+1}^{n} \frac{\cos(2\alpha_i - 2\nu)}{\sigma_i^2}\right] + \Gamma_n \gamma_n^2 = \psi_m + \Delta\psi, \end{aligned}$$

where

$$\Delta\psi = \gamma_m \left[\frac{(1-g_m)\gamma_n}{2} + g_m \sum_{i=m+1}^{n} \frac{\sin^2(\alpha_i - \nu)}{\sigma_i^2}\right] + \Gamma_n \gamma_n^2.$$

Using the above expressions, after some manipulation, the improvement in terms of the localization error due to the increase in the transmit power can be written as

$$\Omega_{\boldsymbol{x}}(P') - \Omega_{\boldsymbol{x}}(P) = \frac{\gamma_m}{\psi_m} - \frac{\gamma_{m+n}}{(1+\chi)\psi_{m+n}} = \frac{\frac{\chi A}{1+\chi} + \frac{B\gamma_n}{(1+\chi)}\left[B+\chi\right] + C}{A\left(A + B\gamma_n\right)} \geq 0, \tag{9.10}$$

where $A = \Gamma_m \gamma_m,\ B = \frac{(1-g_m)}{2}$, and

$$C = g_m \sum_{i=m+1}^{m+n} \frac{\sin^2(\alpha_i - \nu)}{\sigma_i^2} + \frac{\Gamma_n \gamma_n^2}{\gamma_m}.$$

The worst-case geometry for the introduced reference nodes is given by:

$$\alpha_i = \nu, \qquad i = m+1, m+2, \cdots, m+n. \tag{9.11}$$

This implies $\Gamma_n = 0$, and therefore, $C = 0$ in (9.10). Here, due to the lack of knowledge regarding n or γ_n, we assume that the nodes are uniformly distributed with a density ρ and $\beta = 2$. Since $n = N(P') - N(P)$ here depends on χ, assuming n is the average number of new nodes that transmit range response packets after the increase in transmit power, we can write:

$$n = N(P') - N(P) \approx K_N P \chi,$$

where $N(P) = \lfloor \rho \pi R_{\max}^2 \rfloor \approx \rho\pi \left(\frac{\xi_{\min}}{K_P P}\right)^{\frac{2}{\beta}} = K_N P^{\frac{2}{\beta}}$. Additionally, since we do not know the variances σ_i^2, $i = m+1, \cdots, m+n$, we assume they are a factor η larger than σ_m^2. This leads to $\gamma_n = \frac{n}{\eta\sigma_m^2} \approx \frac{K_N P_0 \chi}{\eta\sigma_m^2} = E\chi$, where $E = \frac{K_N P_0}{\eta\sigma_m^2}$. Therefore, substituting $D = \Omega_{\boldsymbol{x}}(P') - \Omega_{\boldsymbol{x}}(P) = \frac{1}{Q_{\boldsymbol{x}}(P')} - \frac{1}{Q_{\boldsymbol{x}}(P)}$ into (9.10), we can solve for χ:

$$\begin{aligned} \chi &= \frac{b \pm \sqrt{b^2 - 4ac}}{2a} \\ a &= BE(DA-1),\ b = (A + B^2E) - DA(A+BE), \\ c &= DA^2. \end{aligned} \tag{9.12}$$

9.9 Appendix 9B: Convergence of Power-Control Algorithms based on Location Estimates

It has been shown that the deterministic scalar power control algorithm

$$p(n+1) = I(p(n))$$

converges to the optimal transmit power $p^\star$ which satisfies $p \geq I(p)$, if the interference function $I(p(n))$ is *standard* [32], [33] in that for $p \geq 0$ it satisfies the following properties:

- Positivity: $I(p) \geq 0$.
- Monotonicity: For $p \geq p'$, $I(p) \geq I(p')$.
- Scalability: For $\alpha \geq 1$, $\alpha I(p) \geq I(\alpha p)$.

It can be shown that if $I(p)$ is standard, the algorithm

$$p(n+1) = (1 - \alpha(n))p(n) + \alpha(n)I(p(n))$$

also converges to $p^\star$, where $0 < \alpha(n) \leq 1$ is the step size at the nth iteration. Consider a power-control algorithm of the form

$$P_{k+1} = I_k(P_k) = P_k + \Delta_P \left(Q_0 - Q(P_k)\right).$$

We have shown in (9.10) that for $p \geq p'$

$$Q(p) \geq Q(p'). \tag{9.13}$$

Therefore, for $p \geq p'$

$$I(p) - I(p') = p - p' + c_Q \left(Q(p') - Q(p)\right) \geq 0,$$

if

$$\Delta_P \leq \frac{p - p'}{Q(p) - Q(p')}, \tag{9.14}$$

$\forall\ p \geq p'$. In order to prove scalability, we note that for $\alpha \geq 1$,

$$Q(\alpha p) \geq \alpha Q(p).$$

Therefore,

$$I(\alpha p) - \alpha I(p) = c_Q(1-\alpha)Q_0 + c_Q \left(\alpha Q(p) - Q(\alpha p)\right). \leq 0$$

We define a stochastic power control algorithm

$$p(n+1) = (1 - \alpha(n))p(n) + \alpha(n)\tilde{I}(p(n), v(n)). \tag{9.15}$$

In the above equation, $\tilde{I}(p(n), v(n))$ is the noisy interference estimate containing the stochastic variables $v(n) = \{v_i(n)\}$, and is called a stochastic interference function. In our case, we have

$$\tilde{Q}(p(n)) = Q(p(n)) + v_Q(n), \tag{9.16}$$

where $v_Q(n) \sim \mathcal{N}(0, \sigma_Q^2(p(n)))$. This is valid for small values of K_R as verified via simulations. Further, from Appendix 9A, $\sigma_Q^2(p(n))$ is an non-increasing function of $p(n)$. Therefore,

$$\begin{aligned} \tilde{I}(p(n), v(n)) &= p(n) + \Delta_P(Q_0 - Q(p(n)) + v_Q(n)) \\ &= I(p(n)) - \Delta_P v_Q(n). \end{aligned}$$

Definition 2. *A stochastic interference function $\tilde{I}(p, \mathbf{v})$ is standard if, for $p \geq 0$, the following conditions are satisfied:*

- **Mean Condition**: The expectation

$$E\left[\tilde{I}(p, v)\right] = I(p), \tag{9.17}$$

  where $I(p)$ is a standard interference function. In this case, since

$$E\left[\tilde{Q}(p(n))\right] = Q(p(n)),$$

  from (9.16), we see that (9.17) is satisfied.

- **Lipshitz Condition** There exists $K_1 > 0$ such that

$$[I(p_1) - I(p_2)]^2 \leq K_1 [p_1 - p_2]^2 .$$

In our case, this implies that

$$\begin{aligned} [I(p_1) - I(p_2)]^2 &= [p_1 - p_2 + \Delta_P (Q(p_2) - Q(p_1))]^2 \\ &\leq (p_1 - p_2)^2 , \end{aligned}$$

from (9.14).

- **Growing Condition** There exists $K_2 > 0$ such that

$$E\left[\left(\tilde{I}(p,v) - I(p)\right)^2\right] \leq K_2(1+p^2).$$

In our case, we have

$$\begin{aligned} &E\left[\left(\tilde{I}(p,v) - I(p)\right)^2\right] = c_Q^2 E\left[v_Q^2\right] \\ &= c_Q^2 \sigma_Q^2(p) \leq K_2(1+p^2), \end{aligned}$$

for some finite K_2. The above relation is valid because $\sigma_Q^2(p)$ is a finite non-increasing function of p, provided an estimate has been computed using the transmit power $p(n)$.

It has been shown in [33] that a stochastic algorithm satisfying these properties converges to $p^\star$ with probability one if the step size $\alpha(n)$ satisfies

$$\sum_{n=0}^{\infty} \alpha(n) = \infty, \ \sum_{n=0}^{\infty} (\alpha(n))^2 < \infty.$$

We see that a stochastic power-control algorithm of the form

$$P_{k+1} = P_k \left(1 - \frac{\alpha}{k}\right) + \frac{\alpha \Delta_P}{k} \left(Q_0 - \hat{Q}(P_k)\right), \tag{9.18}$$

where $0 < \alpha < 1$, converges with probability one. Therefore, a step-size that guarantees convergence is

$$\Delta_{P,k} = \frac{\alpha \Delta_P}{k}, \quad 0 < \alpha < 1.$$

Chapter 10

Conclusions

In Chapter 1, the stated goal of this dissertation was to broadly answer three main questions regarding the design of indoor UWB PoLoNets from the perspective of localization accuracy. In the following, we discuss the answers to these questions in terms of the ideas discussed in the chapters of this dissertation:

Question: *(a) Is it possible to design PoLoNets using a UWB physical layer that satisfy the practical design constraints on localization accuracy?*

Answer: Yes. In Chapter 1, we stated that a practical target localization accuracy in terms of the average localization error for the indoor fire-fighter tracking application discussed was ~ 1 meter^2. It was shown in Chapter 6 through indoor UWB measurements that the application of the proposed localization techniques to range estimates extracted from these measurements resulted in a average root localization error on the order of ~ 0.6 meter^2.

Question: *(b) What are the main factors that assist/impede the achievement of the target localization accuracy?*

Answer: The main factors that impact the achievement of the target localization accuracy in ad hoc UWB PoLoNets are: (i) the accuracy of TOA-based range estimates estimated from UWB signals, (ii) the accuracy of localization techniques and algorithms that utilize these range estimates, (iii) the degradation in range measurement and localization accuracy in NLOS propagation environments, (iv) the propagation of localization error in ad hoc scenarios, (v) the multiple-access scheme applied and the corresponding network and system parameters such as anchor node density, transmit power, mobile speed, etc.

Question: *(c) Can techniques that mitigate these impediments be developed for UWB PoLoNets in order to ensure the desired localization accuracy?*

Answer: The challenges facing the design of UWB PoLoNets for the target localization accuracy relate to the following aspects addressed in the chapters of this dissertation. We briefly summarize the contents in each chapter, along with the salient issues pertaining to

Table 10.1: LOS and NLOS Range Measurement Error Statistics

Range Measurement Error	LOS	NLOS
Mean	0.02 meters	1.90 meters
Standard Deviation of Error	0.23 meters	2.09 meters

UWB PoLoNet design.

- **Range Estimation (Chapter 2):**

  In order to ascertain the degree of localization accuracy that can be provided in UWB PoLoNets, the accuracy of range-estimation possible through the use of UWB signals in different propagation scenarios needed to be quantified. In Chapter 2, we characterized the impact of different parameters on the accuracy of TOA-based range estimation. A simple practical TOA-estimator for ranging using UWB signals was proposed and analyzed. Using such an estimator, the statistics of the range measurement error in LOS and NLOS propagation scenarios were presented, as shown in Table 10.1. Based on measurement results, we modeled the LOS and NLOS range estimates as unbiased and biased Gaussian random variables. We further found that the bias errors in NLOS range estimates can be modeled as exponential random variables. The presence of these bias errors suggested that providing accurate location estimates in indoor NLOS environments is a challenging task, and exacerbates the propagation of localization error in ad hoc scenarios.

- **Location Estimation Algorithms (Chapter 3):**

  Based on the model for UWB range measurements derived in Chapter 2, we studied the accuracy of location estimation in LOS scenarios. The properties of bounds on the accuracy of location estimation were presented, and it was shown that the localization accuracy depends not only the accuracy of range estimates, but also on the geometry of anchors and the number of range estimates. Using a novel generalization of the geometric dilution of precision, we were able to quantify the individual effects of these parameters on the average localization accuracy. Simulation results corresponding to the well-known LS estimator suggested that similar trends are obeyed by practical estimators. We further showed that weighting and iterative refinement of location estimates can lead to considerable gains in practical location estimation, as seen in Table 10.2, at the price of computational complexity.

- **Robust Localization in NLOS propagation environments (Chapter 4):**

  Due to the biased nature of range estimates in NLOS propagation scenarios, the direct use of NLOS range information can result in considerable degradation in localization

Table 10.2: Simulated average root localization error (meters) when $K_E = 0.01$, $L = 10$ meters in LOS scenarios.

Number of LOS Range Estimates (m)	Least-Squares Estimator	Weighted Least-Squares Estimator	Weighted Least-Squares Estimator with Iterative Refinement
3	4.15	4.15	1.56
4	1.06	0.95	0.72
5	0.71	0.61	0.49

accuracy. Thus, the NLOS problem is a main obstacle in obtaining the target localization accuracy in indoor UWB PoLoNets. This necessitates the characterization of the impact of NLOS signal propagation on localization accuracy, and the development of NLOS mitigation algorithms. To this end, we partitioned the NLOS problem into NLOS identification and NLOS mitigation aspects. NLOS identification deals with the identification of NLOS range estimates. We proposed a novel computationally-efficient NLOS identification schemes based on delay spread parameters that can be extracted from the received signal. These methods were found to produce a probability of error of approximately 5%, and further shown to produce negligible degradation in the localization accuracy through measurements.

A novel NLOS mitigation algorithm based on linear-programming (LP) was proposed that relies on NLOS identification. The comparison of the localization accuracy achieved by LP and LS-based estimators with indoor UWB measurements are summarized in Table 10.3. The main advantages of the proposed NLOS mitigation method were: (i) the statistics of the NLOS bias errors are not assumed to be known *a priori*, (ii) no range information is discarded, (iii) this method outperforms the LS estimator in terms of the mean and standard deviation of the localization error, given a mixture of LOS and NLOS range estimates, (iv) this method can be generalized and extended to handle degenerate cases with insufficient (< 3) LOS range estimates, (v) improve localization accuracy in LOS scenarios with a poor geometry of localized nodes, and (vi) can be extended to include other forms of constraints such as physical maps, multi-hop information, GPS data, etc.

- **Propagation of Localization Error (Chapter 5):**

  In the absence of a pre-existing localization infrastructure within the area of interest, ad hoc sequential localization schemes can be used. However, since the location estimates of reference nodes are likely to be noisy, as the distance to the fixed anchors increases, the localization error can propagate leading to poor localization accuracy in regions of the area of interest far away from fixed anchors. We showed that geometry of localized nodes and the bias introduced by practical estimators are key factors that determine

Table 10.3: Average root localization error (meter^2) in NLOS scenarios obtained through measurements.

Root Localization Error (meters)	LS-(LOS+NLOS) Estimator	LS-(Pure LOS) Estimator	LP-Basic Estimator	LP-Extended Estimator
Mean	2.71	1.90	0.53	0.50
Standard Deviation	4.82	4.23	0.70	0.64

Table 10.4: Comparison of CT-SSMA and CSMA in terms of the simulated average root localization error (in meters) values at $t = 20$ seconds after deployment. The parameters used in this simulation can be found in Section 7.4.

Protocol	$L = 10$ meters $\beta = 2$	$L = 10$ meters $\beta = 3$	$L = 20$ meters $\beta = 2$	$L = 20$ meters $\beta = 3$
CT-SSMA	2.74 meters	5.43 meters	2.05 meters	4.70 meters
CSMA	15.51 meters	17.03 meters	16.21 meters	21.91 meters

the extent of the propagation of error. The main implication of the propagation of localization error is that the area (determined by the distance from anchor nodes) over which a desired localization accuracy can be guaranteed may be limited. We showed that this area can be expanded through the use of more efficient (albeit more computationally-intensive) location estimators in LOS scenarios. In NLOS scenarios, the NLOS bias errors were found to increase the propagation of error. A novel multi-hop propagation of error mitigation scheme based on the linear programming method developed in Chapter 4 was proposed. This method incorporates biased NLOS range estimates and mitigates the propagation of localization error. The efficacy of this method was validated through measurements.

- **Multiple-Access Design (Chapters 7 and 8):**

  In Chapter 7, based on insights into the nature of the UWB physical layer, we proposed a spread-spectrum protocol based on time-hopping for UWB PoLoNets, which is based on the common-transmitter spread-spectrum multiple-access approach. Network simulations were performed to demonstrate that the proposed protocol is superior to the CSMA protocol in terms of the localization accuracy and convergence time of the location estimates, as shown in Table 10.4.

  Multiple-access schemes in sensor and mobile ad hoc networks are typically designed from the perspectives of data throughput, latency or energy efficiency. As localization accuracy is the main metric of interest in PoLoNets, the connection between multiple-access protocols and localization accuracy was examined. In Chapter 8, we

demonstrated that the problem of minimizing the average localization error in a single slotted packet radio network was equivalent to maximizing the effective throughput of range estimates, a conjecture that was previously made using insights gained from bounds on location estimation.

The results in Chapter 3 suggested that increasing the number of anchors that provide range estimates improves the average localization accuracy, and the benefits of increasing the number of localized nodes were observed in Chapters 3-6. However, the multiple-access design aspects of successfully obtaining range estimates from a large number of localized nodes was not considered. In Chapter 8, we presented a framework for studying the impact of multiple-access parameters on the average localization accuracy of unlocalized nodes. We showed that in an interference-limited Aloha-based multiple-access scheme, there are optimal values of the system parameters that maximize the average localization accuracy, and these parameters can be extracted if the effective throughput can be extracted as a function of the system parameters. These analytical insights were then verified through simulations. Further, the trade-offs between localization accuracy and the speed of mobile nodes was investigated through analysis and simulations.

- **Power Control Algorithms (Chapter 9):**

  As a mobile node moves through the area of interest, the connectivity with localized reference nodes can fluctuate considerably. Further, if the probability of reference node failure is significant (e.g., fire-fighter tracking network), controlling the transmit power to ensure the minimum required connectivity becomes essential. The main advantage of power-control is: (i) improving connectivity in low-density PoLoNets, thereby ensuring the minimum number range estimates required to compute location estimates, (ii) improving the quality and number of range estimates, thereby improving average localization accuracy. In Chapter 9, we demonstrate that in the absence of multiple-access interference, increase in the transmit power of unlocalized nodes can result in improvement in the location estimates, as shown in Table 10. Two ameliorative power control schemes for a single unlocalized node were discussed: a stochastically convergent iterative scheme based on the quality of location-estimates and a two-step analytical solution. We demonstrated that the two approaches approximate the empirically-obtained optimal transmit power, and significantly improve the robustness of location-estimates. Although the stochastic convergence of the iterative power control algorithm was not proved in the multiple unlocalized node scenario, its efficacy was demonstrated via simulations.

In addition to the above ideas, the results presented in this dissertation provide insights into the overall trade-offs in the design of PoLoNets. Consider the following example: if N_m the number of pulses in a packet that are used to estimate the TOA of the received signal is reduced, then the accuracy of range estimates degrades. The resulting loss in localization accuracy can be compensated in several ways as shown in Figure 10.1. The framework

Table 10.5: Simulated impact of transmit power on the average root localization error (meters) when $K_P = 3.5 \times 10^6$, $K_R = 100$, $N_A = 10$, $L = 10$ meters in LOS scenarios.

Transmit Power P_T (mW)	K_E value	LS Localization Error (meters)
0.5	5.8×10^{-2}	3.19
1	2.9×10^{-2}	2.26
2	1.5×10^{-2}	1.61

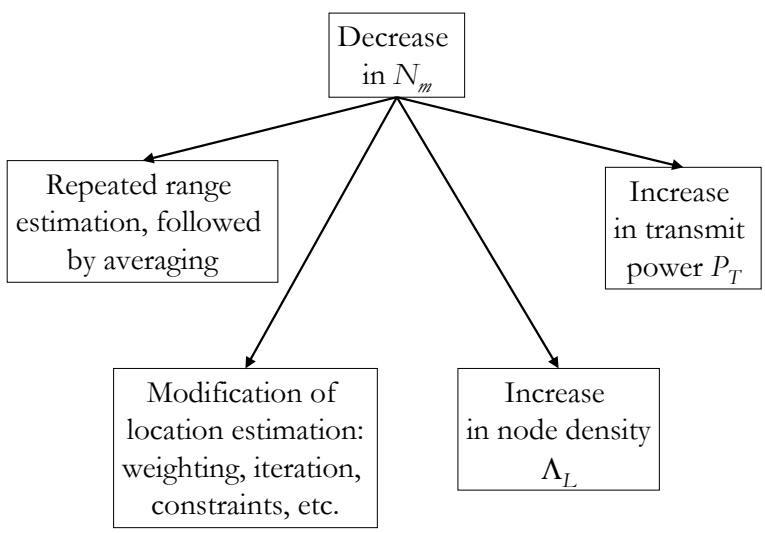


Figure 10.1: Possible design modifications to maintain the localization accuracy in case of a reduction in the packet size.

developed in this dissertation allows us to use a localization budget to modify the parameters in different components of the design. For instance, if the N_m is reduced by a factor of 2, from the analysis of range estimation, the variance of range estimates increases by a factor of two, and so does the average localization error. In order to compensate for this, we know that a 3 dB increase in the transmit power reduces the range estimate variances by a factor of 2, and also the localization error. Similarly, the degradation in localization accuracy can be compensated by trading performance for computational complexity, by using a weighting scheme or iterative position refinement.

In this manner, the insights gained in this dissertation can prove valuable in the system-level design of indoor UWB PoLoNets.

Bibliography

[1] R. Buehrer, W. Davis, A. Safaai-Jazi, and D. Sweeney, "Ultra-wideband Propagation Measurements and Modeling - DARPA NETEX Final Report," tech. rep., Virginia Tech, January 2004. Available online at http://www.mprg.org/ people/buehrer/ultra/ darpa_netex.shtml.

[2] C.-C. Chong, Y. Kim, and S. Lee, "Channel Model Parameterization of the Indoor Residential Environment," tech. rep., IEEE, July 2004. submitted to IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs), available online at http://grouper.ieee.org/ groups/802/15/pub/2004/July04/.

[3] J. Caffery and G. Stuber, "Subscriber Location in CDMA Cellular Systems," *IEEE Transactions on Vehicular Technology*, vol. 47, pp. 406–416, May 1998.

[4] J. Caffery and G. Stuber, "Overview of Radiolocation in CDMA Cellular Systems," *IEEE Communications Magazine*, vol. 36, pp. 38–45, April 1998.

[5] G. Messier and J. Nielsen, "An Analysis of TOA-Based Location for IS-95 Mobiles," in *Proceedings of 1999 Vehicular Technologies Conference*, pp. 1067–1071, 1999.

[6] Y. Shang, W. Ruml, Y. Zhang, and M. Fromherz, "Localization from connectivity in sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 15, pp. 961–974, November 2004.

[7] A. Savvides, H. Park, and M. B. Srivastava, "The n-hop Multilateration Primitive for Node Localization Problems," *Mobile Networks and Applications*, vol. 8, pp. 443–451, 2003.

[8] S. Thrun, W. Burgard, and D. Fox, "A probabilistic approach to concurrent mapping and localization," *Autonomous Robots*, vol. 5, pp. 253–271, 1998.

[9] N. Patwari, A. O. Hero, M. Perkins, N. S. Correal, and R. J. O'Dea, "Relative location estimation in wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2137–2148, Aug. 2003.

[10] C. S. Taggart, Y. Viniotis, and M. L. Sichitiu, "Minimizing average network delay for ultra-wideband wireless networks," in *2005 IEEE Wireless Communications and Networking Conference*, vol. 4, pp. 2276 – 2280, March 2005.

[11] B. Radunovic and J.-Y. Le Boudec, "Optimal power control, scheduling, and routing in UWB networks," *IEEE Journal on Selected Areas in Communications*, vol. 22, pp. 1252 – 1270, September 2004.

[12] R. L. Moses, D. Krishnamurthy, and R. Patterson, "An auto-calibration method for unattended ground sensors," in *ICASSP 2002*, pp. 2941–2944, May 2002.

[13] M. Z. Win and R. A. Scholtz, "Impulse radio: How it works," *IEEE Communications Letters*, vol. 2, pp. 36 – 38, February 1998.

[14] S. J. Ingram, D. Harmer, and M. Quinlan, "Ultra-wideband Indoor Positioning Systems and their Use in Emergencies," in *Position Location and Navigation Symposium, 2004. PLANS 2004*, (Rome, Italy), April 2004.

[15] D. B. Jourdan, D. Dardari, and M. Z. Win, "Position Error Bound and Localization Accuracy Outage in Dense Cluttered Environments," in *2006 IEEE 2006 International Conference on Ultra-Wideband*, pp. 519–524, Sept. 2006.

[16] H. L. V. Trees, *Detection, Estimation, and Modulation Theory.* Wiley, 1968. New York.

[17] J. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Cramer-Rao lower bounds for the time delay estimation of UWB signals," in *2004 IEEE International Conference on Communications*, pp. 3424 – 3428 vol.6, 20-24 June 2004.

[18] J.-Y. Lee and R. A. Scholtz, "Ranging in a dense multipath environment using an UWB radio link," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1677–1683, December 2002.

[19] C. Chang and A. Sahai, "Estimation bounds for localization," in *2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks, (IEEE SECON 2004)*, pp. 415 – 424, 4-7 Oct. 2004.

[20] J. J. Caffery, "A new approach to the geometry of TOA location," in *2000 IEEE Vehicular Technology Conference*, vol. 4, pp. 1943 – 1949, September 2000.

[21] K. Yu and I. Oppermann, " Performance of UWB position estimation based on time-of-arrival measurements," in *2004 International Workshop on Ultra Wideband Systems*, pp. 400 – 404, May 2004.

[22] S. Venkatesh and R. M. Buehrer, "Power Control in UWB Position-Location Networks," in *2006 IEEE International Conference on Communications*, vol. 9, pp. 3953 – 3959, June 2006.

[23] S. Venkatesh and R. M. Buehrer, "Multiple-access design for ad hoc UWB position-location networks," in *2006 IEEE Wireless Communications and Networking Conference*, vol. 4, pp. 1866 – 1873, June 2006.

[24] T. Silventoinen and M. Rantalainen, "Mobile station emergency locating in GSM," in *1996 IEEE International Conference on Personal Wireless Communications*, pp. 232 –238, Feb. 19-21 1996.

[25] K. Yu, "3-D Localization Error Analysis in Wireless Sensor Networks," *submitted to IEEE Transactions on Wireless Communications*, 2006.

[26] S. Venkatesh and R. M. Buehrer, "NLOS Identification in Ultra-Wideband Systems based on Received Signal Statistics," *To appear in IEE Proceeding on Microwaves, Antennas and Propagation: Special Issue on Antenna Systems and Propagation for Future Wireless Communications*, March 2006.

[27] S. Venkatesh and R. M. Buehrer, "A Linear Programming approach to NLOS error mitigation in Sensor Networks," in *Proceedings of the Fifth International Conference on Information Processing in Sensor Networks (IPSN '06)*, vol. 1, pp. 301 – 308, April 2006.

[28] S. Venkatesh and R. M. Buehrer, "NLOS Mitigation using Linear Programming in Ultra-Wideband Location-Aware Networks," *to appear in IEEE transactions on Vehicular Technology*, February 2006.

[29] M. Hellebrandt and R. Mathar, " Location tracking of mobiles in cellular radio networks," *IEEE Transactions on Vehicular Technology*, vol. 48, pp. 1558 – 1562, September 1999.

[30] E. S. Sousa and J. A. Silvester, "Spreading code protocols for distributed spread-spectrum packet radio networks," *IEEE Transactions on Communications*, vol. 36, pp. 272–281, February 1988.

[31] E. S. Sousa and J. A. Silvester, "Optimum transmission ranges in a direct-sequence spread-spectrum multihop packet radio network," *IEEE Journal on Selected Areas in Communications*, vol. 8, pp. 762 – 771, June 1990.

[32] R. D. Yates, "A framework for uplink power control in cellular radio systems," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1341 – 1347, Sept. 1995.

[33] J. Luo, S. Ulukus, and A. Ephremides, "Standard and quasi-standard stochastic power control algorithms," *IEEE Transactions on Information Theory*, vol. 51, pp. 2612 – 2624, July 2005.

[34] S. M. Kay, *Fundamentals of Statistical Processing, Volume I : Estimation Theory.* Prentice-Hall Inc., 1993. 2nd Edition.

[35] M. Z. Win and R. A. Scholtz, "Ultra-Wide Bandwidth Time-Hopping Spread-Spectrum Impulse Radio for Wireless Multiple-Access Communications," *IEEE Transactions on Communications*, vol. 48, April 2000.

[36] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios," *IEEE Signal Processing Magazine*, vol. 22, pp. 70– 84, July 2005.

[37] V. S. Somayazulu, "Multiple access performance in UWB systems using time hopping vs. direct sequence spreading," in *2002 IEEE Wireless Communications and Networking Conference*, vol. 2, pp. 522–525, March 2002.

[38] B. B. Denis, J. B. Pierrot, and C. Abou-Rjeily, "Joint distributed synchronization and positioning in uwb ad hoc networks using toa," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 4, pp. 1896 – 1911, 2006.

[39] I. Guvenc and Z. Sahinoglu, "Threshold selection for UWB TOA estimation based on kurtosis analysis," *IEEE Communication Letters*, vol. 9, pp. 1025–1027, December 2005.

[40] D. Dardari and M. Z. Win, "Threshold-based Time-of-arrival Estimators in UWB Dense Multipath Channels," in *2006 IEEE International Conference on Communications*, vol. 10, (Istanbul), pp. 4723–4728, June 2006.

[41] C. Falsi, D. Dardari, L. Mucchi, and M. Z. Win, "Time of Arrival Estimation for UWB Localizers in Realistic Environments," *EURASIP Journal on Applied Signal Processing*, vol. 1, pp. 1–13, June 2006.

[42] M. Abramovitz and I. A. Stegun, *Handbook of mathematical functions.* Dover, 1970.

[43] B. Denis, J. Keignart, and N. Daniele, "Impact of NLOS propagation upon ranging precision in UWB systems," in *2003 IEEE Conference on Ultra Wideband Systems and Technologies*, pp. 379 – 383, November 2003.

[44] B. Alavi and K. Pahlavan, " Bandwidth effect on distance error modeling for indoor geolocation," in *2003 IEEE Proceedings on Personal, Indoor and Mobile Radio Communications (PIMRC 2003)*, vol. 3, pp. 2198 – 2202, September 2004.

[45] M. A. Spirito, "On the accuracy of cellular mobile station location estimation," *IEEE Transactions on Vehicular Technology*, vol. 50, pp. 674–685, May 2001.

[46] D. J. Torrieri, "Statistical theory of passive location systems," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 20, no. 2, pp. 183 – 198, 1984.

[47] K. Yu and I. Oppermann, "UWB positioning for wireless embedded networks," in *2004 IEEE Radio and Wireless Conference, (RAWCON 2004)*, vol. 1, pp. 459 – 462, September 2004.

[48] R. J. Fontana, E. Richley, and J. Barney, "Commercialization of an ultra wideband precision asset location system," in *2003 IEEE Conference on Ultra Wideband Systems and Technologies, (UWBST 2003)*, vol. 1, pp. 369 – 373, November 2003.

[49] M. P. Wylie and J. Holtzman, "The Non-Line of Sight Problem in Mobile Location Estimation," in *5th IEEE International Conference on Universal Personal Communications*, vol. 2, pp. 827 – 831, Sept. 1996.

[50] J. Borras, P. Hatrack, and N. B. Mandayam, "Decision theoretic framework for NLOS identification," in *1998 Vehicular Technology Conference, (VTC '98)*, vol. 2, pp. 1583 – 1587, May 18-21 1998.

[51] S. S. Ghassemzadeh, R. Jana, C. W. Rice, W. Turin, and V. Tarokh, "Measurement and modeling of an ultra-wide bandwidth indoor channel," *IEEE Transactions on Communications*, vol. 52, pp. 1786 – 1796, October 2004.

[52] Y. Qi and H. Kobayashi, "Cramer-Rao lower bound for geolocation in non-line-of-sight environment," in *Proceedings of ICASSP 2002*, vol. 3, pp. III–2473 – III–2476, May 2001.

[53] Y. Qi, H. Kobayashi, and H. Suda, "Analysis of Wireless Geolocation in a Non-Line-of-Sight Environment," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 672 – 681, March 2006.

[54] L. Doherty, K. pister, and L. El Ghaoui, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, (Anchorage, AK), pp. 1655–1663, Apr. 22–26, 2001.

[55] X. Wang, Z. Wang, and R. O'Dea, "A TOA-based location algorithm reducing the errors due to non-line-of-sight (NLOS) propagation," *IEEE Transactions on Vehicular Technology*, vol. 52, pp. 112–116, January 2003.

[56] E. G. Larsson, "Cramer-Rao bound analysis of distributed positioning in sensor networks," *IEEE Signal Processing Letters*, vol. 11, pp. 334–337, March 2004.

[57] P.-C. Chen, "A non-line-of-sight error mitigation algorithm in location estimation," in *Wireless Communications and Networking Conference 1999*, vol. 1, pp. 316–320, Sept. 21-24th 1999.

[58] S. Al-Jazzar, J. Caffery, and H. R. You, "A scattering model based approach to NLOS mitigation in TOA location systems," in *2002 IEEE Vehicular Technology Conference (VTC-Spring 2002)*, vol. 2, p. 861865, May 2002.

[59] Z. Irahhauten, H. Nikookar, and G. J. Janssen, "An Overview of Ultra Wide Band Indoor Channel Measurements and Modeling," *IEEE Microwave and Wireless Components Letters*, vol. 14, pp. 386 – 388, August 2004.

[60] J.-M. Cramer, M. Z. Win, and R. A. Scholtz, "Evaluation of the Multipath Characteristics of the Impulse Radio Channel," in *The Ninth IEEE International Symposium on Personal Indoor and Mobile Radio Communications*, vol. 2, pp. 864–868, 1998.

[61] S. M. Yano, "Investigating the Ultra-wideband Indoor Wireless Channel," in *IEEE VTS 55th Vehicular Technology Conference*, vol. 3, pp. 1200–1204, October 2002.

[62] S. S. Ghassemzadeh, L. J. Greenstein, T. Sveinsson, A. Kavcic, and V. Tarokh, "UWB Delay Profile Models for Residential and Commercial Indoor Environments," *IEEE Transactions on Vehicular Technology*, vol. 54, pp. 1235 – 1244, July 2005.

[63] J. Foerster, "Channel Modeling Sub-committee Report Final (Doc: IEEE 802-15-02 / 490r1-SG3a)," tech. rep., IEEE, February 2002. Available online at http:// grouper.ieee.org/ groups/802/15/pub/2002/Nov02/.

[64] S. S. Ghassemzadeh, L. J. Greenstein, A. Kavcic, T. Sveinsson, and V. Tarokh, "UWB indoor delay profile model for residential and commercial environments," in *2003 IEEE Vehicular Technology Conference, (VTC 2003-Fall)*, vol. 5, pp. 3120 – 3125, 6-9 October 2003.

[65] C.-C. Chong, Y.-E. Kim, S. K. Yong, and S.-S. Lee, "Statistical characterization of the UWB propagation channel in indoor residential environment," *IEEE Transactions on Communications*, vol. 5, pp. 503 – 512, August 2005.

[66] A. F. Molisch, "IEEE 802.15.4a channel model-final report," tech. rep., IEEE, September 2004. (IEEE 802.15-04-0662-01-04a, Sept. 2004).

[67] Y. Qi and H. Kobayashi, "On relation among time delay and signal strength based geolocation methods," in *Proceedings of IEEE Global Telecommunications Conference (GLOBECOM '03)*, vol. 7, pp. 4079–4083, December 2003.

[68] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, *Linear Programming and Network Flows.* John Wiley and Sons, New York, 1995.

[69] B. Denis, L. Ouvry, B. Uguen, and F. Tchoffo-Talom, " Advanced Bayesian filtering techniques for UWB tracking systems in indoor environments," in *2005 IEEE International Conference on Ultra-Wideband (ICU 2005)*, vol. 1, pp. 638–643, September 2005.

[70] C. Wann, Y. Chen, and M. Lee, "Mobile location tracking with NLOS error mitigation," in *2002 IEEE Global Telecommunications Conference (GLOBECOM '02)*, vol. 2, pp. 1688 – 1692, November 2002.

[71] J. M. Huerta and J. Vidal, "Mobile tracking using UKF, time measures and LOS-NLOS expert knowledge," in *IEEE ICASSP '05*, vol. 4, pp. iv/901 – iv/904, March 2005.

[72] C. Savarese, J. Rabaey, and J. Beutel, "Location in distributed ad-hoc wireless sensor networks," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol. 4, (Salt Lake City, UT), pp. 2037–2040, May 7–11, 2001.

[73] A. Savvides, C.-C. Han, and M. B. Srivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proceedings of the 7th annual international conference on Mobile computing and networking*, (Rome, Italy), pp. 166 – 179, 2001.

[74] M. Rabbat and R. Nowak, "Decentralized source localization and tracking," in *Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04). IEEE International Conference on*, vol. 3, pp. 921–4, May 2004.

[75] S. Basagni, M. Battelli, M. Iachizzi, C. Petrioli, and M. Salehi, "Limiting the propagation of localization errors in multi-hop wireless network," in *Pervasive Computing and Communications Workshops, 2006. PerCom Workshops 2006. Fourth Annual IEEE International Conference on*, Mar. 13–17, 2006.

[76] D. Niculescu and B. Nath, "Error characteristics of ad hoc positioning systems (APS)," in *Proceedings of the 5th ACM international symposium on Mobile ad hoc networking and computing*, (Tokyo, Japan), pp. 20–30, May 2004.

[77] R. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 1985.

[78] J. A. Costa, N. Patwari, and A. O. Hero, "Distributed weighted-multidimensional scaling for node localization in sensor networks," *ACM Transactions on Sensor Networks (TOSN)*, vol. 2, no. 1, pp. 39 – 64, 2006.

[79] D. B. Jourdan, J. Deyst, J. J., M. Z. Win, and N. Roy, "Monte Carlo localization in dense multipath environments using UWB ranging," in *2005 IEEE International Conference on Ultra-Wideband*, pp. 314–319, Sept. 5–8, 2005.

[80] D. Bertsekas and R. Gallagher, *Data Networks*. Prentice Hall, 1992. 2nd Edition.

[81] R. Jurdak, C. Lopes, and P. Baldi, "A Survey, Classification and Comparative Analysis of Medium Access Control Protocols for Ad Hoc Networks," *IEEE Communications Surveys and Tutorials*, vol. 6, pp. 2–16, Q1 2004.

[82] I. Oppermann, L. Stoica, A. Rabbachin, Z. Shelby, and J. Haapola, "UWB wireless sensor networks: UWEN - a practical example," *IEEE Communications Magazine*, vol. 42, pp. S27–S32, December 2004.

[83] A. Papoulis, *Probability, Random Variables and Stochastic Processes*. McGraw-Hill Inc., 1991. 3rd Edition.

[84] J. Ilow and D. Hatzinakos, "Analytic alpha-stable noise modeling in a Poisson field of interferers or scatterers," *Signal Processing, IEEE Transactions on [see also Acoustics, Speech, and Signal Processing, IEEE Transactions on]*, vol. 46, pp. 1601–1611, June 1998.

[85] W. Feller, *A Introduction to Probability Theory and its Applications*. Wiley, 1966. Vol. II.

Vita

Swaroop Venkatesh was born in Bangalore, India in 1980. His early schooling was at the Kendriya Vidyalaya, Indian Institute of Science, Bangalore. After finishing high school in 1998, he joined the Department of Electrical Engineering at the Indian Institute of Technology (IIT) Madras. As part of the dual degree program, he received the degrees of Bachelor and Master of Technology (B. Tech and M. Tech) in Electrical Engineering from IIT Madras in 2003.

In 2003, Venkatesh began work towards a Ph.D. degree in Electrical Engineering with the Mobile and Portable Radio Research Group at Virginia Polytechnic Institute and State University. His research included ultra-wideband communication and localization, wireless propagation studies, and communication system design. These research activities were supplemented by a summer internship at Time Domain Corporation, Huntsville, Alabama, in 2004. He received his Ph.D. degree in Electrical Engineering from Virginia Polytechnic Institute and State University in May 2007.