UNIVERSITY OF CALIFORNIA
Santa Barbara

Imaging Sensor Nets:

Scalable architectures for data collection and localization in sensor networks

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

by

Bharath Ananthasubramaniam

Committee in Charge:

Professor Upamanyu Madhow, Chair

Professor Jerry D. Gibson

Professor João P. Hespanha

Professor Gregory J. Pottie

Professor Mark J. Rodwell

Professor Kenneth Rose

December 2007

UMI Number: 3291293

The Dissertation of
Bharath Ananthasubramaniam is approved:

Professor Jerry D. Gibson

Professor João P. Hespanha

Professor Gregory J. Pottie

Professor Mark J. Rodwell

Professor Kenneth Rose

Professor Upamanyu Madhow, Committee Chairperson

September 2007

Imaging Sensor Nets:

Scalable architectures for data collection and localization in sensor networks

by

Bharath Ananthasubramaniam

To my grandmother

Acknowledgements

I am deeply grateful to my parents, without whom I literally would not be here, and my sister Aparna, for being everything I am not. I dedicate this dissertation to my grandmother, whose home improvement assignments initiated me into science and engineering. I am in debt of all my teachers, who laid the foundations for this work. My undergrad years would have been fruitless if not for some of my closest friends - Ananth, Anil, Avinash, Jagadish, Kapil, MV, Paro Ramji, Shiva and Vinay - and our collaborative efforts at learning Electrical Engineering.

I am extremely thankful to my advisor, Prof. Upamanyu Madhow, for his guidance and wisdom through my doctoral work. I thank my committee for their insights and helpful suggestions to improve this work. I was also blessed with a warm and friendly environment in my lab: Gwen, for her help and encouragement in those early years, when research is hardest. Raghu, lent his unique perspective on many rewarding discussions. Munkyo, for the long nights of experimenting with the prototype. Kris, Noah, Sumit, Jaspreet, Eric, Sriram and Sandeep, for the memorable times at work.

I am much obliged to these beacons of support who helped me sail through these last five years of highs and lows so far from home: Chethan, Gautham, Ashwin and Kostas. Finally, the people who added color to this journey through Santa Barbara : Anup, Anuroop, Belinda, Steffen, Natalie, Petko, Tobias, Lucile, and David.

Curriculum Vitæ

Bharath Ananthasubramaniam

Education

2007	Ph.D. Electrical Engineering,
	University of California, Santa Barbara
	Advisor: Prof. Upamanyu Madhow
2004	M.S. Electrical Engineering,
	University of California,Santa Barbara
2002	B.Tech. Electrical Engineering,
	Indian Institute of Technology, Madras, India

Journal Publications

B. Ananthasubramaniam and U. Madhow, "On Localization Performance in Imaging Sensor Nets", *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 5044-5057, Oct. 2007.

Conference Publications

M. Seo, B. Ananthasubramaniam, U. Madhow and M. Rodwell, "Millimeterwave Imaging Sensor Nets: A Scalable 60-GHz Wireless Sensor Network", Proc. *International Microwave Symposium (IMS2007)*, Honolulu, Hawaii, Jun. 2007.

B. Ananthasubramaniam and U. Madhow, "Collector receiver design for data collection and localization in sensor-driven networks", Proc. *Conference on Information Sciences and Systems (CISS'07)*, Baltimore, MD, Mar. 2007.

E. Torkildson, B. Ananthasubramaniam, U. Madhow and M. Rodwell, "Millimeter-wave MIMO: Wireless Links at Optical Speeds", (Invited Paper) Proc. of 44th *Allerton Conference on Communication, Control and Computing*, Monticello, Illinois, Sept. 2006.

B. Ananthasubramaniam and U. Madhow, "Detection and Localization of Events in Imaging Sensor Nets", Proc. 2005 *IEEE International Symposium on Information Theory (ISIT)*, Adelaide, Australia, Sept. 2005

B. Ananthasubramaniam and U. Madhow, "Virtual Radar Approach to Event Localization in Sensor Networks", Proc. 2004 *IEEE International Symposium on Information Theory (ISIT)*, Chicago, IL, Jun. 2004

B. Ananthasubramaniam and U. Madhow, "Virtual Radar Imaging for Sensor Networks", Proc. *Third International Symposium on Information Processing in Sensor Networks (IPSN'04)*, Berkeley, CA, Apr. 2004

B. Ananthasubramaniam, G. Barriac, R. Mudumbai and U. Madhow, "Distributed space-time communication for sensor networks", First *International Symposium on Control, Communications and Signal Processing*, 2004.

Abstract

Imaging Sensor Nets:
Scalable architectures for data collection and localization in sensor networks

Bharath Ananthasubramaniam

We present *imaging sensor nets* that provide scalable architectures by moving the complexity from the sensor, which plays the role of pixels, to the collectors. This analogy to imaging addresses the key issues of data collection and localization in large-scale randomly deployed sensor networks; in such networks, conventional multihop relay to extract the data is often inapplicable and sensor geolocation capability is expensive. The sensors can be made extremely "dumb" and low-cost with minimum functionality (no geolocation and networking). In one instance of an *imaging sensor net*, a sophisticated airborne collector node queries the sensor field with a radio frequency beacon. Sensors electronically reflect the beacon adding low-rate data, thus, creating a virtual radar geometry, which provides fine-grained localization, and conveying data to the collector. We show that these reflecting sensors can be located using a Synthetic Aperture Radar-like two dimensional matched filtering algorithm and its decision-directed extensions.

This concept is translated into a millimeter-wave prototype, where a stationary collector sweeps a sensor field with a spread-spectrum beacon that is reflected by an RFID-like sensor after imposing data on it. The collector is built from

off-the-shelf components and the *semi-passive* sensors are implemented on low-cost printed circuit boards. We develop low-complexity algorithms for collector baseband processing to demodulate the data and locate the sensor. Preliminary experimental results demonstrate the feasibility of our approach and ability to support data rates up to 100 kbps, while providing a few centimeters location accuracy.

We investigate source localization based on angle of arrival (AoA) measurements at a geographically dispersed network of collectors in real-world propagation environments. Accordingly, a sequential localization algorithm capable of suppressing the outlying AoA measurements due to multipath is presented. This algorithm achieves close to optimal performance at a complexity that is linear in the number of measurements. A possible application of this algorithm is in another instance of an *imaging sensor net*, where the sensor transmits its data without prior coordination with a network of collectors that are responsible for locating the sensor and decoding the data.

Contents

Acknowledgements **v**

Curriculum Vitæ **vi**

Abstract **ix**

List of Figures **xiii**

1 Introduction **1**
1.1 Dissertation Overview 6
1.2 Literature Review and Related Work 8
1.3 Organization 13

2 SAR-like Localization in Imaging Sensor Nets **14**
2.1 System Model 16
2.1.1 Received Signal Model 17
2.1.2 Spatial Representation for an Airborne Collector 21
2.1.3 Spatial Representation for a Stationary Collector 23
2.2 Imaging Algorithms for Sensor Localization 25
2.2.1 Two Dimensional Matched Filtering 26
2.2.2 Maximum Likelihood Single Sensor Localization 27
2.2.3 Decision-directed Localization of Multiple Sensors 29
2.2.4 Standard SAR Algorithm versus ML Algorithm 31
2.3 Noise-limited Localization Performance 35
2.3.1 Scale-Invariant System Dimensions 35
2.3.2 Range Resolution 36
2.3.3 Azimuth Resolution 38
2.3.4 Running Example for Simulation Model 43
2.3.5 Trade-offs between Parameters 44

2.4 Performance with inter-sensor interference 51
2.4.1 Termination Criterion for Decision-directed Algorithm . . 54
2.4.2 Decision-directed algorithm versus Optimal ML Localization 57
2.4.3 Simulation Results for Dense Sensor Deployment 59
2.5 Conclusions 62

3 Signal Processing for a Millimeter-wave Prototype 64
3.1 The Millimeter-wave Prototype 66
3.1.1 Radio Link Characterization for Semi-passive and Active Sensors 68
3.1.2 System Parameter Choices 70
3.1.3 Prototype Collector 71
3.1.4 Prototype Sensor 75
3.2 Collector Baseband Processing 77
3.2.1 Received Signal Model 78
3.2.2 Data Extraction 85
3.2.3 Locating the Sensor in Range 88
3.3 Experiments on the Prototype 91
3.3.1 Experimental Methodology 91
3.3.2 Results 93
3.4 Discussion of Issues and Possible Fixes 98
3.5 Conclusions 102

4 Angle-of-Arrival-based Localization using a Network of Collectors 104
4.1 Models of AoA Estimates 107
4.2 Localization in LOS scenarios 111
4.2.1 LMMSE Updates & Transformations 112
4.2.2 Sequential Localization Algorithm 116
4.2.3 ML Estimate and Cramer-Rao Bound 119
4.2.4 Properties of the Sequential Algorithm 122
4.3 Localization in NLOS settings 124
4.3.1 ML Localization 125
4.3.2 Sequential Aggregation with Outlier Suppression 126
4.4 Numerical Results 132
4.4.1 Performance under LOS scenarios 133
4.4.2 Performance under NLOS scenarios 136
4.5 Conclusions 145

5 Conclusions **148**
5.1 Summary 148
5.2 Challenges and Future Work 150

Bibliography **155**

Appendix **164**

A Bounds on bootstrap failure probability **165**

List of Figures

1.1 Examples of Imaging Sensor Net architectures 4

2.1 Imaging sensor net with a moving collector 18
2.2 A collector-driven system with a stationary collector. 24
2.3 Active sensor locations in standard SAR and ML localization comparison. 32
2.4 Output of standard SAR processing for active sensors in Figure 2.3 33
2.5 Output of ML algorithm for active sensors in Figure 2.3 34
2.6 Comparison between standard SAR and ML techniques 34
2.7 Random variables and errors in azimuth processing. 39
2.8 Effect of SNR on range estimate 45
2.9 Effect of SNR on azimuth estimate 46
2.10 Effect of signal bandwidth on range estimate 48
2.11 Effect of signal bandwidth on azimuth estimate 49
2.12 Effect of antenna beamwidth on range estimate 50
2.13 Effect of antenna beamwidth on azimuth estimate 51
2.14 Effect of antenna beam shape on range estimate 52
2.15 Effect of antenna beam shape on azimuth estimate 53
2.16 Receiver Operational Characteristic 56
2.17 Decision-directed versus ML algorithm with two active sensors . . 59
2.18 Performance of decision-directed algorithm for 10 sensors 60
2.19 Performance of decision-directed algorithm for 50 sensors 61

3.1 Block diagram of the 60 GHz prototype collector 72
3.2 A photograph of the collector brassboard setup. 74
3.3 A prototype 60-GHz low-cost passive sensor 76
3.4 Block diagram of the Signal Processing Algorithm 81
3.5 Prototype BER performance at 100 kbps 95

3.6 Verification of Radio Link Characterization 96
3.7 Variation of estimated range with sensor collector distance 97
3.8 RMS localization error at different sensor locations 98

4.1 Sensor field with 8 collectors on its perimeter 113
4.2 Sequential localization performance in LOS scenarios 133
4.3 Sequential localization performance in LOS scenarios with random bootstraps 134
4.4 Convergence of the sequential algorithm to the ML estimate . . . 135
4.5 Dependence of localization performance on number of collectors . 136
4.6 Performance of NLOS suppression with optimal thresholds 138
4.7 Estimation cost for NLOS suppression 139
4.8 Performance of NLOS suppression with realistic threshold choice . 140
4.9 Observed probability of localization failure 141
4.10 NLOS suppression performance under alternate multipath models 142
4.11 Ray-tracing illustration in narrow multipath setting 144
4.12 Ray-tracing illustration in wideband multipath setting 146

Chapter 1

Introduction

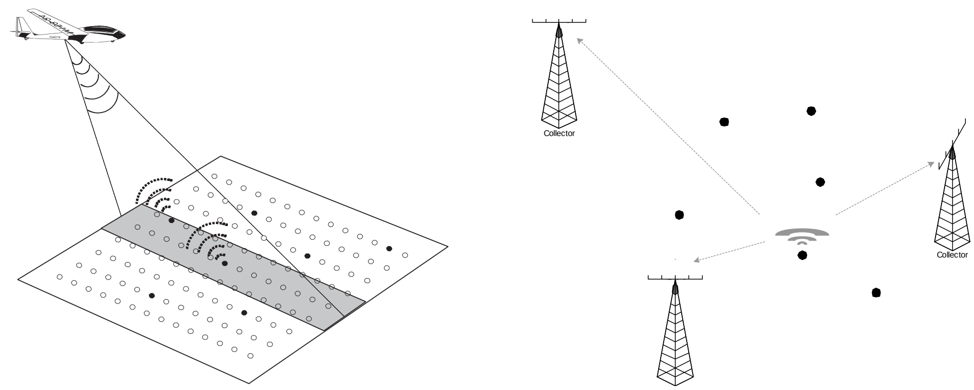
> *The whole is greater than the sum of its parts. The part is greater than its role in the whole.*
>
> Tom Atlee

A sensor network is a collection of devices with limited sensing, computational and communication capability that collaborate to observe the environment in which they are placed[1]. Such networks find application in a wide variety of areas such as habitat monitoring[2], meteorological and geophysical studies[3], inventory management, smart homes[4], surveillance, vehicle tracking, early alert systems[5] for fire or earthquake emergencies, to name only a few. The appeal of these networks lies in their ability to pervasively sense and monitor desired phenomena across a range of spatial and temporal scales using an assortment of sensing modalities[6]. These networks are a natural outcome of the recent explosion in wireless devices, along with the trend of miniaturization and functional integration. Moreover, as the ability to measure or sense is an essential component of most engineering systems, there are abundant future applications for sensor networks.

Very large-scale sensor networks arise naturally when low-cost microsensors with small sensing range are used to provide coverage of a large area. Such networks have a number of important applications, including border policing, monitoring for biological or chemical agents over large urban areas, and even interplanetary exploration. The ability to randomly deploy large numbers (hundred to thousands) of sensors and operate them remotely (e.g., deployment and subsequent monitoring from either aircraft or spacecraft) is key to their utility and poses important challenges to the design of these networks. The extraction of data from the network of sensors is a significant bottleneck, since communication is the most energy-consuming sensor operation. Energy is at a premium, as the sensors are expected to operate self-sufficiently for an extended period of time after deployment, without intervention. Since the locations where the measurements were gathered is an integral part of the data and the locations of randomly-deployed sensors cannot be known a priori, localization of the sensors becomes vital.

The conventional approach of multihop forwarding of data through the sensor network to a collection node is often ineffective in scenarios where the collector node is remotely located. For example, in applications in which a distant sensor field is monitored by an aircraft, low-earth orbit satellites or stationary monitoring facilities, multihop relay does not significantly reduce distance to the collector. Furthermore, multihop networking does not, in general, scale well with the size of the network as demonstrated by the pioneering work of Gupta and Kumar[7] and subsequent analyses of the capacity of wireless networks[8, 9, 10, 11], with the throughput per node going to zero as the network size increases. In such multihop networks, sensors need to be equipped with a networking stack and run protocols for inter-sensor networking, while incurring the additional overhead for addressing and routing. Moreover, multihop communication might not always be the most energy-efficient method of data extraction from the sensor network[12]. Traditional localization solutions such as GPS[13] are expensive or might be unavailable due to jamming, and algorithms for sensor self-localization require added functionality such as signal strength indicators or multiple antennas to compute inter-sensor distances.

In this dissertation, we propose novel architectures to simultaneously address the requirements of scale, localization and data collection in such randomly deployed large-scale sensor networks. To this end, we draw inspiration from imaging (ranging from passive optical to active radar) to develop a class of *Imaging Sensor Nets*, interpreting sensors as pixels imaged by a sophisticated collector node. The number of "pixels" in the image can be scaled to millions and is limited only by the signal-to-noise ratio (SNR) and capability (e.g., aperture size) of the collector. Although conventional imaging is confined to phenomena with characteristic electromagnetic signatures, sensors can be used as transducers to translate arbitrary phenomena into data than can be recovered using radio-frequency (RF) techniques. Further, the location of a "pixel" in an image of the sensor field is easily recovered, given knowledge of the location and orientation of the sophisticated imaging node.



(a) Collector-driven System (b) Sensor-driven System

Figure 1.1: Examples of Imaging Sensor Net architectures

We now present two complementary instances of *Imaging Sensor Nets* that are akin to synthetic aperture radar (SAR)[14] imaging and radio interferometry[15], respectively:

1. An airborne collector initiates communication by illuminating a part of the sensor field with an RF beacon, as shown in Figure 1.1(a). Sensors illuminated by the beacon and with data to send electronically reflect the beacon and modulate it with local data, resulting in a *virtual* radar geometry. Multiple overlapping illuminations of the sensor field called 'snapshots' are jointly processed by the collector using radar and imaging techniques to achieve fine-grained localization of the reflecting sensor, and multiuser demodulation algorithms are used to recover the sensor data. Although a single collector is sufficient, localization and demodulation gains can be had with multiple collectors. A key feature of this architecture is that the data

collection is *collector-driven*, having serious implications on data storage and representation, and latency in data recovery.

2. The *sensor-driven* paradigm, on the other hand, offers a complementary approach to data collection. In a *sensor-driven* network (see Figure 1.1(b)), the sensor transmits data as soon as it observes an 'interesting' event without prior coordination with a geographically dispersed network of collectors that constantly monitor for sensor transmissions. The onus is on the network of collectors to cooperatively detect, demodulate and locate the sensor transmission. Due to the absence of a virtual radar geometry, multiple collectors are necessary to locate the sensor and are not just a performance enhancing feature as in the *collector-driven* system.

These asymmetric system architectures allow drastic reduction in the sensor node functionality by moving the complexity to the collector nodes. The sensors can be "dumb" without incurring additional functional overhead for inter-sensor networking and geolocation. Since this functional reduction helps pare down the data sent by the sensor to a bare minimum, link-budgets to communicate over long distances (upto 200 km for a satellite-based collector) are feasible even with severely energy constrained nodes.

1.1 Dissertation Overview

In this dissertation, we restrict our attention to the design of the communication link between the sensors and collector(s) for data collection and localization in these large-scale imaging sensor nets. The detailed design of sensing systems that utilize these architectures are not addressed here, but some key issues involved are highlighted in the conclusions (Section 5).

We first focus on the fundamental localization capabilities of a *collector-driven* network monitored by an airborne collector. Consequently, we consider an idealized sensor communication model, where the sensor transmits one bit using on-off signaling, which is applicable even if the sensor modulates low-rate data. The sensors are either active and reflect the collector's beacon, or inactive and do not. The standard SAR algorithm applied to this system performs poorly due to the lack of phase synchronization between the collector and sensor.

Hence, we develop a ML localization algorithm for a single active sensor and show that it is a modification of standard SAR processing. ML localization reduces to two-dimensional matched filtering across time and across multiple snapshots. Since the ML estimator for multiple active sensors is computationally intractable, we resort to a decision-directed approach to sequentially locate all the active sensors. The noise-limited performance of these algorithms compares well with the Cramer Rao bound and the localization resolution scales with the SNR at the receiver, the bandwidth of the beacon, and the spatial sampling rate of the collector. In dense deployments with inter-sensor interference, the ability to locate at least one sensor is each cluster of active sensors is demonstrated.

We, next, describe a millimeter-wave prototype based on this *collector-driven* architecture, in which a stationary collector sweeps the sensor field with a mechanically steered antenna. The collector transmits a periodic spread-spectrum (SS) location code that is reflected by a *semi-passive* sensor after low-rate modulation and frequency translation to avoid backscatter. The collector is built using off-the-shelf components consisting of a brassboard collector and low-cost printed circuit board (PCB) *semi-passive* sensors (similar to Radio Frequency identification (RFID) tags[16]). This prototype is currently under development in collaboration with our colleagues Munkyo Seo and his advisor Prof. Mark Rodwell in electronics at UCSB.

Using the insights gained from the idealized model, we develop a low-complexity algorithm for processing the prototype collector baseband outputs to demodulate the sensor data and locate the sensor. We report on preliminary experiments characterizing the data demodulation and range estimation performance of the prototype, and demonstrate the capability of this architecture to provide good localization resolutions while sustaining data rates of several kbps. We also discuss sources of observed losses in performance and suggest possible solutions.

Finally, we investigate cooperative localization of a source using a network of geographically dispersed receivers using AoA. Source localization using a network of receivers has numerous applications including localization of the transmitting sensor in a *sensor-driven* network, which was the initial motivation for this work. We present a sequential algorithm for updating source location estimates in the presence of line-of-sight (LOS) between the source and the receivers. However, multipath scattering and reflections are often encountered in practical propagation environments and this non-line-of-sight (NLOS) multipath often results in *outliers* in the AoA measurements. Using the sequential algorithm as a building block, we develop an algorithm that estimates the source location while suppressing the effect of these *outliers*. ML localization in such scenarios requires exhaustive testing of estimates from all possible subsets of measurements. We avoid this by utilizing a randomized algorithm that approaches the ML performance at a complexity that is only linear in the number of measurements. The localization error is proportional to the AoA error variance and search area, and can be reduced by increasing the number of receivers. Our results also show that the capability of the receiver to resolve multiple incoming paths is vital to achieving good performance in NLOS settings.

1.2 Literature Review and Related Work

We now review some of the important developments on data collection and localization recognizing that a rich history of sensor network literature on this topic exists. This is followed by a discussion of systems very similar to those proposed in this dissertation.

As mentioned earlier, multihop relay is the method of choice for data collection in wireless sensor networks and traces its origins to early developments in wireless ad hoc networks. Much of the research in data collection has focussed on optimizing multihop networking in the specific context of sensor networks. These include application-specific routing [17, 18, 19, 20], medium access control (MAC) layer optimization [21, 22], topology control [23, 24], quality of service [25, 26], distributed source coding and aggregation[27, 28, 29, 30], and distributed signal processing and estimation [31, 32].

Despite the limits on scalability derived by Gupta and Kumar [7] and others, it is conceivable that in certain settings, the redundancy in the information gathered by the sensor nodes may be such that the net information to be conveyed to a data collection center scales up slowly enough to fit within the Gupta-Kumar bounds [33, 34]. However, as noted earlier, even if the problems of scale and overhead in multihop networking could be circumvented, it is inapplicable to a large class of applications served by imaging sensor nets, namely, those in which the nodes in the sensor field are more or less at the same distance from the collector node, and all nodes have comparable energy/power constraints.

For more conventional multihop architectures, there has been a great deal of activity in the important problem of localization[1], broadly classified into two categories. The majority of schemes fall into the first category of *anchor-based* localization [35, 36, 37, 38, 39, 40], in which a subset of the sensors know their locations, and the information from their beacons is used by other sensor nodes to infer their own locations in an iterative fashion. In the second category of *anchor-free* localization [41], the nodes compute their locations iteratively in a consistent coordinate system. It must be noted that in contrast to the "dumb" nodes in an imaging sensor net, all of the preceding methods require some form of ranging (at different degrees of sophistication) in the sensor nodes, followed by distributed collaborative iterations.

Related Work: The proposed *collector-driven* paradigm is closely related to the work of the SAR research group at Sandia National Labs [42], Hounam et al.[43] and Colpitts et al.[44]. These *passive* RF tags that respond to queries from a radar transmitter are used for identification and localization in [42, 43] and for insect tracking in [44]. These tags are restricted to respond to a radar transmission by modulating a fixed data or identification sequence, as the tags draw all the needed power from the radar signal. On the other hand, in the proposed *semi-passive* sensors, the addition of a power source at the sensor provides low-rate data modulation capability.

Our prototype is similar, in principle, to the work of Stoleru et al.[45] and the global positioning system (GPS)[13]. Sensors are coarsely localized by a collector using a spot beam in [45]. However, our objective is to obtain an accurate estimate of the sensor locations without stringent requirements on the collector's beams. In GPS, low-rate navigational data modulated on a coarse/acquisition SS code is transmitted by satellites, and the GPS receiver on earth locates itself using differential range measurements from the SS code in conjunction with the navigational data. There are, however, some key differences between the proposed prototype design and GPS. In the imaging sensor net prototype, the collector transmits the SS signal and then locates the reflecting sensor transceiver without the sensor computing or knowing its own position. Moreover, unlike GPS only a single collector is necessary to locate the sensor, although multiple collectors could be used to augment performance. Further, we aim for larger bit-rates of the order of 100s of kbps as against the 50 bps in GPS, while suffering from residual frequency and timing offsets due to the data modulation and location code being imposed on the signal at different locations.

An example of a *sensor-driven* network is the Remote Battlefield Sensor System (REMBASS) and its improvement (IREMBASS)[46] used by the US Department of Defence for surveillance and situational awareness. Although REMBASS networks consist of transmit-only sensors, the sensor locations are determined during deployment, thus, limiting the size and speed of deployments. Locating a transmitting source using a network of receivers is a historically well studied source localization problem with several sensor network based applications such as bird habitat monitoring using bird calls and surveillance using camera networks. Source localization algorithms have been proposed for utilizing a variety of measurement modalities such as RSS [47, 48, 49] and TDOA [50, 51], despite the unreliability of RSS and the difficulty in achieving synchronization between receivers. However, acoustic systems can capitalize on better receivers and slower propagation speeds of sound, and TDOA has been used in underwater [52] and sensor network applications[53, 54, 55]. However, all these algorithms require LOS channels between the source and the collectors with no multipath. In practice, errors in TDOA, AoA or RSS due to multipath scattering effects dominate performance[56].

As explained earlier, the outliers produced by NLOS multipath must be disregarded before the emitter can be reliably located. The two approaches common in the estimation literature [57] involve either identification and removal of the *outliers* before location estimation or robust estimation of the position while simultaneously limiting the effect of outliers. Using the former approach, outliers are eliminated using statistics of the NLOS estimates [58, 59, 60] and probabilistic combining of estimates [61, 62]. We adopt, however, the latter robust estimation approach that is most similar to the work of Chen [63] and Casas et al.[64]. Chen[63] proposes a weighted least-squares localization using range measurements, where the weights are iteratively varied to assign the lowest weights to outliers and highest to the LOS estimates, but suffers from exponential complexity in the number of measurements. Casas et al. [64] solve the multilateration by a least median squares algorithm on time-of-flight (TOF) estimates, which simultaneously eliminates outliers and estimates the emitter location. Although this least median squares approach produces the "best" robust estimate of the source location using three "good" TOF measurements, the algorithm does not utilize all the "good" TOF measurements to enhance the localization. On the contrary, we propose a robust estimation-based approach that has linear complexity in the number of measurements and tries to estimate the sensor location using the maximal set of non-outlying measurements.

1.3 Organization

This dissertation is organized as follows. An instance of a *collector-driven* imaging sensor net with an airborne collector is described and localization algorithms for such a system are developed in Chapter 2. The hardware design and baseband signal processing algorithms for a prototype imaging sensor net with a stationary collector are presented in Chapter 3. In Chapter 4, we develop low-complexity algorithms for cooperative localization of a source using AoA measurements that are also capable of suppressing outliers due to NLOS multipath. Finally, concluding remarks and future directions of research are provided in Chapter 5.

Chapter 2

SAR-like Localization in Imaging Sensor Nets

A picture is worth a thousand words.

Chinese proverb

In this chapter, we elaborate on the *collector-driven* architecture with an airborne collector for scalable data collection and localization in very large-scale sensor networks. Since, accurate localization of the sensors that detect an event, or have data to report, is an essential feature of an imaging sensor net, we focus on developing fundamental insight into the localization performance achievable in an imaging sensor net, using an idealized model which ignores data modulation. This is a good approximation when low-rate data modulation is imposed on, say, a spread spectrum beacon being reflected by the sensor, as is the case in the imaging sensor net prototype presented in Chapter 3. Indeed, the insights gained from this idealized model are used to design the prototype, with the localization algorithms developed here serving as fundamental blocks in the collector baseband processing. This work was presented previously in [65] and [66].

Under this idealized model, sensors respond using one-bit, on-off keying to the airborne collector's beacon. Sensors are either "active" or "inactive", with only the active sensors electronically reflecting the beacon, thereby creating a geometry as in Synthetic Aperture Radar (SAR) [67, 68, 14]. However, two-dimensional matched filtering used in standard SAR processing performs poorly because of the lack of carrier synchronization between the sensor nodes and the collector as shown in Section 2.2.4.

Our main results are as follows.

1. We provide a Maximum Likelihood (ML) formulation for localization, considering first an isolated active sensor, which leads to a noncoherent decision statistic based on a simple modification of the two-dimensional matched filter. The model is similar to that in noncoherent radar tomography [69, 70]. Our ML algorithm also applies to multiple active sensors, provided that they are spaced far enough apart that their two-dimensional responses at the collector do not overlap.

2. Since ML localization for multiple active sensors that are closely spaced is computationally intractable, we develop a suboptimal decision feedback algorithm, in which the estimated response of each active sensor is subtracted out once it is detected. A criterion for terminating the algorithm is provided, based on an analysis of the probabilities of false alarm and miss.

3. Key tradeoffs governing localization performance are investigated analytically. Simulation results are provided, and compared with analysis when applicable.

The chapter is organized in the following manner. The system model is presented in Section 2.1. The optimal ML single sensor localization algorithm and a decision-directed joint localization of multiple sensor are developed in Sections 2.2.2 and 2.2.3. Noise-limited localization performance, assuming a single active sensor, is considered in Section 2.3. This is used to gain insight into the effect of appropriately chosen dimensionless parameters on performance measures. The performance with multiple active sensors is explored in Section 2.4. Finally, Section 2.5 contains concluding remarks.

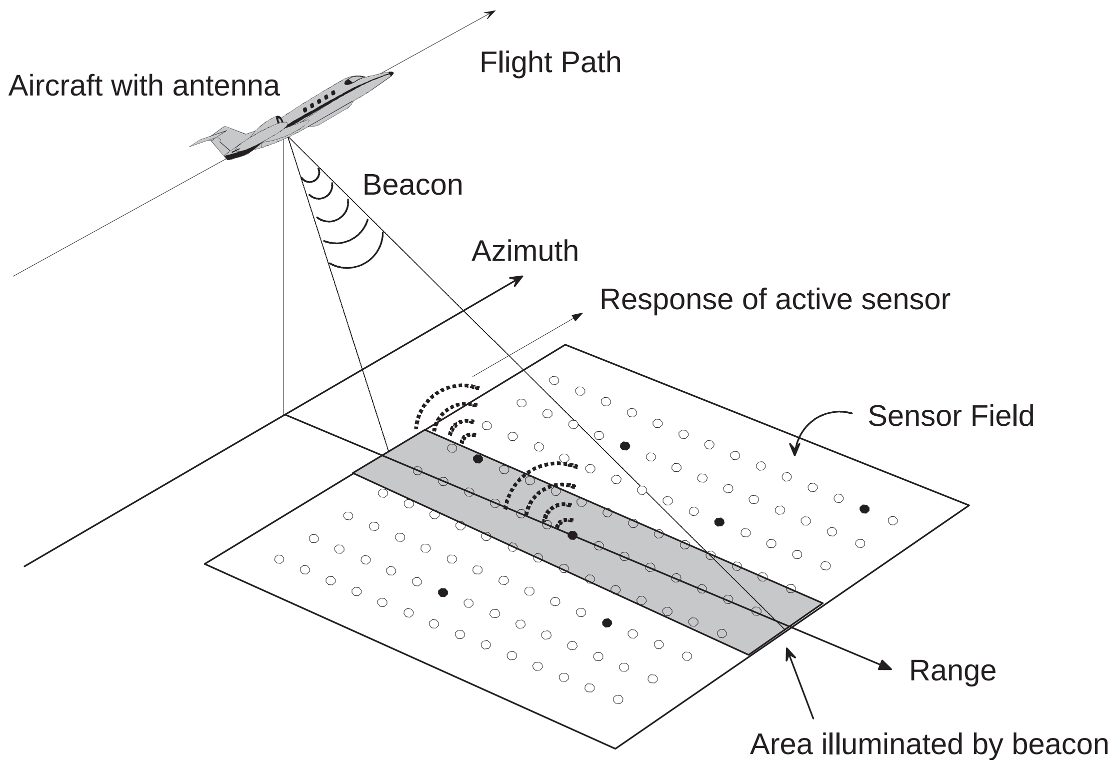
2.1 System Model

In this section, we describe the system model corresponding to an airborne collector (e.g., an aircraft or UAV), in direct analogy to swath-mode SAR, as shown in Figures 2.1(a) and 2.1(b). However, these concepts extend directly to other geometries, such as terrestrial vehicles moving along the edge of the sensor field, or stationary collectors with steered beams as illustrated in Section 2.1.3. The collector node illuminates a part of the field with a beacon using a side-looking antenna. Each such illumination is called a *snapshot*. As seen in Figure 2.1(b), the collector moves along one edge of the sensor field at a fixed altitude, and the movement of the collector causes the beacon to sweep the entire field.

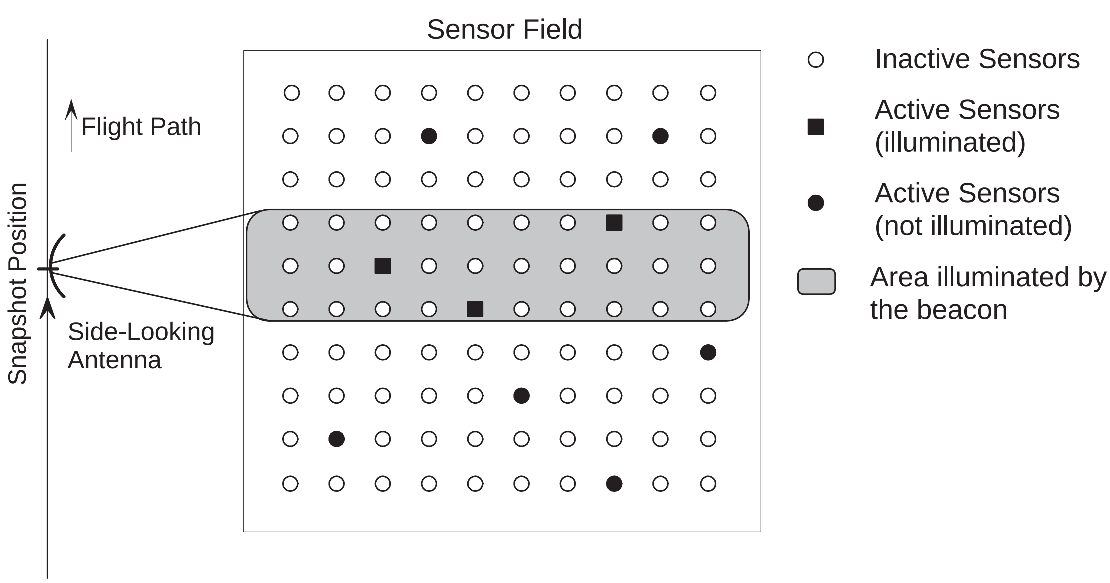
Ignoring sensor data modulation, we obtain an idealized one-bit model of the sensor data. Sensors are either "active" or "inactive," and the objective is to localize the active sensors (i.e., to image the activity in the sensor field). Active sensors that hear the beacon respond to it by transmitting a wideband signal, timing their response precisely with respect to a trigger sequence in the beacon. This creates a SAR-like geometry. The collector node processes the net received signal over multiple snapshots using SAR-like [67][68][14] and noncoherent tomography based [70] techniques to generate an image of the activity in the sensor field. The collector node knows its own location at the time of different snapshots (e.g., an aircraft may know its own GPS location, and its height relative to the sensor field). Thus, the collector can estimate the absolute locations of the active sensors, up to the resolution of this virtual radar system.

2.1.1 Received Signal Model

Each active sensor sends back a complex baseband signal, $s(t)$, modulated on a sinusoidal carrier of frequency f_0. Nevertheless, the techniques presented here can be easily extended to settings where each active sensor sends back a different signal (e.g., waveforms randomly chosen from a near-orthogonal set) in order to mitigate inter-sensor interference. The transmitted passband signal is $\tilde{s}(t) = Re\{s(t)e^{j2\pi f_0 t}\}$. Suppose, there are K active sensors, indexed by k, on the



(a) An aircraft collecting data from a sensor field



(b) Top view of the virtual radar system

Figure 2.1: Imaging sensor net with a moving collector

field, and the collector takes J snapshots, indexed by j, of the field. Note that the time reference for each snapshot is different: at each snapshot, at the instant the collector's beacon is transmitted, the time variable is reset to zero to simplify subsequent notation. The complex baseband received signal, $r_j(t)$, at the collector node at snapshot j is

$$r_j(t) = \sum_{k=1}^{K} h_{j,k}\ \tilde{I}_{j,k}\ s(t - \tau_{j,k})e^{-j2\pi f_0 \tau_{j,k}} + n_j(t), \quad j = 1, \ldots, J,$$

where $h_{j,k}$ is a complex channel gain, $\tau_{j,k}$ is the round-trip propagation time between active sensor k and the collector node at snapshot j, $\tilde{I}_{j,k}$ is the antenna gain function (AGF) of the collector antenna, which is the antenna gain to sensor k in snapshot j, and $n_j(t)$ is the noise. Note that $r_j(t)$ will be a vector if the collector has multiple receive antennas.

The channel gain, $h_{j,k}$, captures the effects of multipath fading, signal path loss, and lack of synchronization between the local oscillators (LOs) at the sensors and the collector. We assume an Additive White Gaussian Noise (AWGN) channel with line of sight (LOS) communication. The path loss in the signal is ignored for two reasons: (i) the path loss exponent can vary significantly (between 2 and 6) depending on factors such as atmospheric conditions, aperture-medium coupling, frequency of operation, and (ii) incorporating this exponent into the estimator does not provide significant improvement in performance, as, in practice, estimates of the exponent are coarse. Since the path loss is ignored, the received signal-to-noise ratio (SNR) is the same for all sensors. Nevertheless, to account for path loss, this SNR can be replaced by the minimum SNR, seen by the farthest sensor, to obtain a conservative estimate of performance. (Fading and shadowing effects, if any, can be accommodated by an outage analysis not undertaken here.)

The LOs at the sensors and the collector are not synchronized. However, the frequency offset between the oscillators is assumed to be small enough that the relative phase is constant over the duration of the transmitted pulse. The relative phases from one snapshot to another are modeled as independent and identically distributed (i.i.d.) over $[0, 2\pi]$. Note that this is a worst case scenario, where no attempt is made to track the frequency drift and phase offset of the sensor LO. Tracking the frequency and phase offsets of the sensor LO provides improved performance at the cost of more computation at the collector. Under the preceding assumptions, the complex gains for this noncoherent AWGN LOS channel are

$$h_{j,k} = e^{j\theta_{j,k}} \quad k = 1, ..., K, \quad j = 1, ..., J \ ,$$

where $\theta_{j,k}$ are i.i.d. and uniform over $[0, 2\pi]$. Absorbing all deterministic phases into the random phase factor, the received signal reduces to,

$$r_j(t) = \sum_{k=1}^{K} \ \tilde{I}_{j,k} \ s(t - \tau_{j,k}) e^{j\theta_{j,k}} + n_j(t), \ j = 1, \ldots, J \ . \tag{2.1}$$

The round-trip delay $\tau_{j,k} = \frac{2R_{j,k}}{c}$, where $R_{j,k}$ is the distance between the collector node and sensor k in snapshot j, and c is the speed of light. This is identical to conventional radar: the start transmission field reaches active sensor node k at a time $\frac{R_{j,k}}{c}$ after it is generated by the collector node, and the response of sensor k reaches the collector node at a time $\frac{R_{j,k}}{c}$ after it is transmitted by active sensor k.

2.1.2 Spatial Representation for an Airborne Collector

The received signal model in (2.1) is used to develop localization algorithms in Section 2.2. It is sometimes convenient to use a spatial received signal representation, to view the sensor field as an 'image' or to analyze performance. The y-axis of the coordinate system is chosen to be the airborne collector's flight path. The locations of the K active sensors are $\{(x_k, y_k)\}_{k=1}^{K}$, and the location of the collector at snapshot j is $(0, z_j = jD)$, where the distance between snapshots is D. The round-trip delay between the sensor and the collector in snapshot j is

$$\tau_{j,k} = \frac{2R_{j,k}}{c} = \frac{2\sqrt{x_k^2 + (y_k - z_j)^2}}{c} = \frac{2x_k}{c}\sqrt{1 + \frac{(y_k - z_j)^2}{x_k^2}}.$$

The AGF $\tilde{I}_{j,k} = I(z_j - y_k)$ and the beamwidth of the antenna, $2B$, is such that

$$I(y) \leq \epsilon \quad \forall\ |y| > B\ , \tag{2.2}$$

for some chosen antenna gain ϵ. Generally, ϵ can be chosen to include the first side-lobes or a significant part of the main lobe. However without any loss of generality in the algorithms in Section 2.2, we assume an antenna with only a main lobe. In practice, an antenna is characterized by an angular beamwidth Θ_{BW}(defined analogously with B in (2.2)) and B varies with distance from the antenna, R, as $B = R\Theta_{BW}$. We consider, here, two AGFs: (a) an ideal beam which has unity gain within the beamwidth. Therefore, the AGF $\tilde{I}_{j,k} = 1$ when sensor k is active and is illuminated by the beacon in snapshot j, and 0 otherwise. (b) A more realistic Gaussian beam approximation of a parabolic antenna,

$$I(y) = \frac{\sqrt{2}}{\sqrt{\pi}B_{3dB}} e^{\frac{-2y^2}{B_{3dB}^2}},\ |y| < B,$$

where B_{3dB} is the half-power beamwidth of the antenna ($\epsilon = 0.5$) and $B = B_{3dB}$. The number of times an active sensor is illuminated by the collector, N, is the ratio of the antenna beamwidth and the distance between snapshots,

$$N = \frac{2B}{D}.$$

By the standard far-field approximation employed in SAR, and assuming a highly directional antenna, $x_k \gg B$, the round-trip delay can be approximated as

$$\tau_{j,k} \approx \frac{2x_k}{c} = \tau_k \ , \tag{2.3}$$

observing that only active sensors illuminated by the beacon, satisfying $|y_k - z_j| \leq B$, respond. In other words, a sensor's distance from the collector node is approximated by the 'x' coordinate of the sensor, and sensors at the same range in a given snapshot lie approximately on a line parallel to the y-axis. The error due to this approximation is

$$\tau_{j,k} - \tau_k \approx \frac{(y_k - z_j)^2}{cx_k} \ \leq \ \frac{B^2}{cx_k}, \tag{2.4}$$

since $|y_k - z_j| \leq B$.

Each sensor is now associated with a single delay τ_k. Hence, (2.1) becomes

$$r_j(t) = \sum_{k=1}^{K} \tilde{I}_{j,k} \ s(t - \tau_k) e^{j\theta_{j,k}} + n_j(t), \quad j = 1, ..., J. \tag{2.5}$$

By a coordinate transformation that maps the time-coordinate at each snapshot (recall the time variable is reset at each snapshot) to the 'x' coordinate, and the location of the collector node z_j to the 'y' coordinate, i.e.,

$$t = \frac{2x}{c} \ \text{ and } \ z_j = y \ , \tag{2.6}$$

equation (2.5) becomes

$$r(x, y) = \sum_{k=1}^{K} I(y - y_k) s \left(\frac{2(x - x_k)}{c} \right) e^{j\theta_k(y)} + n(x, y) . \tag{2.7}$$

Note also that the snapshot index j is incorporated into the variable y, i.e., $y = \{jD : j = 1, ..., J\}$. In practice, the received signal is sampled leading to a received signal matrix, whose rows and columns represent discrete values of 'x' and 'y' respectively. This provides an elegant way of mapping the entire sensor field to a matrix and to visualize the output of the processing as an image. The received signal representations in (t, j) and (x, y) are equivalent, and either (2.5) or (2.7) is used in Section 2.2.

2.1.3 Spatial Representation for a Stationary Collector

In this section, we demonstrate the similarity between the spatial representation derived in (2.7) and corresponding representation for a system with a stationary collector. Indeed, this similarity will be exploited to design signal processing algorithms for the prototype in Chapter 3.

A stationary collector sweeps the sensor field by steering the beacon as shown in Figure 2.2 and illuminations along different directions generate multiple snapshots. In other words, an angular sweep produces an effect equivalent to the movement of the airborne collector. Hence, the model in (2.1) can be utilized with the understanding that the snapshot index j now refers to the orientation of the

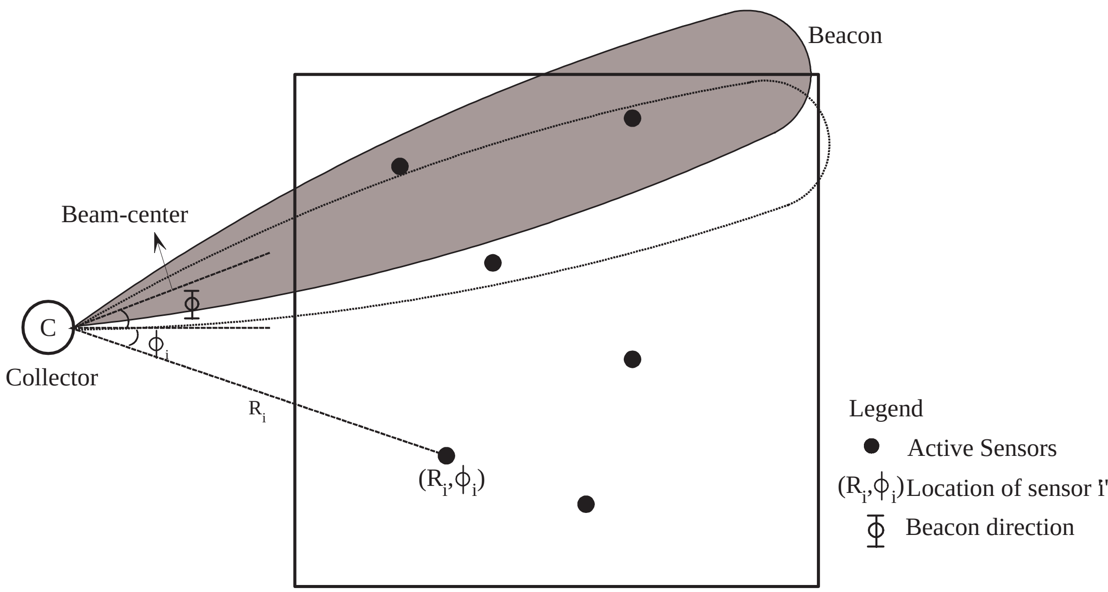


Figure 2.2: A collector-driven system with a stationary collector.

stationary collector's antenna and $\tilde{I}_{j,k}$ is the angular AGF:

$$r_j(t) = \sum_{k=1}^{K} \tilde{I}_{j,k}\ s(t - \tau_{j,k})e^{j\theta_{j,k}} + n_j(t),\ j = 1, \ldots, J\ .$$

We now define the location of the K sensors as $\{(R_k, \phi_k)\}_{k=1}^{K}$ in polar coordinates centered at the collector, the choice of which will become shortly become obvious. Since the distance between the collector and sensor remains constant, $\tau_{j,k} = \tau_k = 2R_k/c$ without the need for the standard far field approximation in (2.3) and we arrive at a model,

$$r_j(t) = \sum_{k=1}^{K} \tilde{I}_{j,k}\ s\left(t - \frac{2R_k}{c}\right) e^{j\theta_{j,k}} + n_j(t), \quad j = 1, ..., J.$$

analogous to (2.5). If the orientation of the collector antenna is Φ then $\tilde{I}_{j,k} = I(\Phi - \phi_k)$. Mapping the time-coordinate at each snapshot to a range coordinate, $t = \frac{2R}{c}$, as in (2.6) the spatial representation for the stationary collector is obtained:

$$r(R, \Phi) = \sum_{k=1}^{K} I(\Phi - \phi_k)\ s\left(\frac{2(R - R_k)}{c}\right) e^{j\theta_k(\Phi)} + n(R, \Phi). \tag{2.8}$$

We observe from (2.7) and (2.8) that the received signal representations for the airborne and stationary collector vary only in their choice of coordinate framework. This implies that the localization algorithms presented in this chapter are all directly applicable to the stationary collector setting as well as long as the received signal represented appropriately.

2.2 Imaging Algorithms for Sensor Localization

Although the virtual radar system geometry is analogous to that of SAR, the standard SAR reconstruction algorithm (presented in Section 2.2.1) for sensor localization is inapplicable due to the lack of coherence between the local oscillators at the sensors and the collector. The effect of this lack of coherence on the SAR algorithm is shown in Section 2.2.4. In Section 2.2.2, we present an ML formulation of the problem, assuming a single active sensor, which leads to a noncoherent decision statistic. The ML single sensor algorithm is then used in Section 2.2.3 as a building block for a suboptimal decision feedback algorithm for localizing multiple active sensors.

2.2.1 Two Dimensional Matched Filtering

We now develop the standard SAR processing for the virtual radar model in (2.7). In standard SAR, due to true reflection at the target, there is phase coherence (only deterministic phases due to propagation delay remain) and (2.7) reduces to

$$\begin{aligned} r(x,y) &= \sum_{k=1}^{K} I(y-y_k)s\left(\frac{2(x-x_k)}{c}\right)e^{-jkx_k} + n(x,y) \\ &= \rho(x,y) \star h(x,y) + n(x,y)\,, \end{aligned}$$

where

$$h(x,y) = I(y)s\left(\frac{2x}{c}\right) \quad \text{and} \quad \rho(x,y) = \sum_{k=1}^{K} \delta(x-x_k, y-y_k)e^{-jkx_k}.$$

with $\star$ representing convolution. ρ represents the locations of the active sensors along with the deterministic phases, and h is the impulse response of this virtual radar system. Recovering $\rho(x,y)$ involves deconvolution or inversion of the filter $h(x,y)$. Thus, the received signal is matched filtered against $h(x,y)$, i.e. $r(x,y)$ is convolved with $h^*(-x,-y)$,

$$\hat{\rho}(x,y) = \rho(x,y) \star h^*(-x,-y). \tag{2.9}$$

Since h is separable in the x and y coordinates, the two-dimensional (2-D) matched filtering can be decomposed into two sequential one-dimensional filtering operations. In the first step, the received signal is processed using a filter matched to the transmitted signal, $s(\frac{2x}{c})$. The output of the first step is passed through a filter matched to the AGF, $I(y)$. These operations along the x and y directions are commonly termed *range correlation* and *azimuth correlation* respectively in the SAR literature.

When this processing is used in the virtual radar system, it does not perform well, as seen from the simulations in Section 2.2.4. This performance degradation due to the unaccounted random phase terms, $e^{j\theta_{j,k}}$, can be eliminated by an ML formulation that accounts for these terms, which is developed now in Section 2.2.2.

2.2.2 Maximum Likelihood Single Sensor Localization

From (2.5), the received signal in snapshot j due to a single active sensor at $\underline{X} = (x_0, y_0) \equiv (\tau, \{\tilde{I}_j\}_{j=1}^{J})$ is

$$r_j(t) = \tilde{I}_j \, s(t-\tau) \, e^{j\theta_j} + n_j(t), j = 1, ..., J, \tag{2.10}$$

where θ_j are i.i.d. uniform random variables, $\tilde{I}_j = I(jD - y_0)$ is the antenna gain to the sensor in snapshot j, and $\tau = \frac{2x_0}{c}$ is the propagation delay between the sensor and the collector in snapshot j. The received signal vector is

$$\mathbf{r} = \mathbf{s} + \mathbf{n} \,, \tag{2.11}$$

where

$$\mathbf{r} = \begin{pmatrix} r_1(t) \\ \vdots \\ r_J(t) \end{pmatrix} \quad , \qquad s_j = \tilde{I}_j \, s(t-\tau) \; e^{j\theta_j},$$

$\mathbf{n}$ is an AWGN vector, and $\mathbf{s}$ is the active sensor response in the absence of noise.

The location estimate is obtained by maximizing the log-likelihood function, or equivalently, minimizing the Euclidean distance between the received signal and transmitted signal vectors jointly over $\underline{X}$ and θ_j,

$$\begin{aligned} \hat{\underline{X}}_{ML}(y) &= \arg\min_{\underline{X},\{\theta_j\}} \|\mathbf{r} - \mathbf{s}(\underline{X})\|^2 \\ &= \arg\min_{\underline{X},\{\theta_j\}} \sum_{j=1}^{J} \|r_j - \tilde{I}_j\, s(t-\tau)\, e^{j\theta_j}\|^2 \\ &= \arg\max_{\underline{X},\{\theta_j\}} \sum_{j=1}^{J} \tilde{I}_j\ |\langle r_j, s(t-\tau)\rangle| \cos(\theta_j - \phi_j), \\ &\qquad\qquad \phi_j = \angle\{\langle r_j, s(t-\tau)\rangle\}\,. \end{aligned} \tag{2.12}$$

To maximize (2.12), the $\cos(\theta_j)$'s are replaced by their maximum likelihood estimates, i.e., $\theta_j = \phi_j$. The simplified likelihood function is

$$\hat{\underline{X}} = \arg\max_{\underline{X}} \sum_{j=1}^{J} \tilde{I}_j\ |\langle r_j, s(t-\tau)\rangle|\,. \tag{2.13}$$

The decision statistic is $\sum_{j=1}^{J} \tilde{I}_j\ |\langle r_j, s(t-\tau)\rangle|$, and the sufficient statistic is $\langle r_j, s(t-\tau)\rangle$.

In (2.12), $\tilde{I}_j = I(jD - y_0)$ is independent of x_0. However in reality, $\tilde{I}_j$ is a function of x_0 too, due to the broadening of the beam with distance from the antenna described in Section 2.1.2. This dependence does not change the form of (2.12), and (2.13) can be modified following the same procedure as

$$\hat{\underline{X}} = \arg\max_{\underline{X}} \sum_{j=1}^{J} \tilde{I}_j(\tau)\ |\langle r_j, s(t-\tau)\rangle| - E_s \sum_{j=1}^{J} \tilde{I}_j(\tau)^2\,. \tag{2.14}$$

where $\tilde{I}_j(\tau) = \tilde{I}(jD - y_0, x_0)$ and $E_s = \langle s, s\rangle$. We infer from (2.14) that the filter used for the azimuth correlation must be varied (lengthened) as a function of the delay τ, which makes the processing more computationally intensive but maintains the simple 2-D filtering structure.

Although a filter matched to $s(t)$ produces the sufficient statistics as in SAR, the optimal processing is nonlinear. The unknown phases, $\{\theta_j\}$, cause the optimal processing to be *noncoherent,* using only the magnitudes of the range correlation for the azimuth processing. These magnitudes are processed using a filter matched to the AGF defined in (2.2). Thus, a minor modification of the standard SAR algorithm produces the ML-estimation rule for a single active sensor localization. Moreover, if the active sensors are sparsely distributed in the sensor field, then there is no interference between active sensor transmissions. Consequently, the multi-event localization can be performed by repeated application of the single sensor algorithm until all the sensors are detected. A suboptimal algorithm to perform multiple sensor localization, in the presence of interference, is presented in Section 2.2.3.

2.2.3 Decision-directed Localization of Multiple Sensors

The received signal from K active sensors in snapshot j is

$$r_j = \sum_{k=1}^{K} \tilde{I}_{j,k} s(t-\tau_k) e^{j\theta_{j,k}} + n_j(t),\ j = 1, ..., J, \tag{2.15}$$

where $\{\theta_{j,k}\}$ is the set of all unknown random phases at the active sensors. When the active sensors are sparsely distributed on the sensor field with no inter-sensor interference, the single sensor algorithm in Section 2.2.2 is optimal for multi-sensor ML localization. Two active sensors interfere with each other when their transmissions are overlapping in the 2-D received signal space consisting of one temporal and one spatial dimension (see Section 2.1.1). In the presence of inter-sensor interference, the optimal ML joint localization algorithm is computationally intractable. Hence, we now present a suboptimal joint detection and localization algorithm that trades off optimality for a lower computational cost.

We adopt a sequential zero-forcing decision-feedback approach where, when an active sensor is detected, it is localized, and the influence of its transmission on the received signal is estimated and subtracted out. This updated received signal is then used to detect (and localize) the next active sensor. The number of active sensors, K, is not known *a priori,* hence this process is continued until a termination criterion is met.

The sequential detection algorithm is initiated by assuming a single active sensor in the field. The ML localization algorithm for a single sensor is used to obtain an estimate of that sensor's location as

$$\hat{X}_1 = (\hat{\tau}_1, \{\hat{I}_{j,1}\}) = \arg\max_X \sum_{j=1}^{J} \tilde{I}_j \; |\langle r_j, s(t-\tau)\rangle| \, , \tag{2.16}$$

where the estimates, $\hat{\tau}$ and $\{\hat{I}_{j,1}\}$, uniquely determine the location of the sensor. Assuming the estimate $\hat{X}_1$ is the true location, the effect of this active sensor on the received signal is subtracted out. The response of this active sensor also depends on the phases, $\{\theta_{j,1}\}$, for which ML estimates were obtained in (2.12) as

$$\hat{\theta}_j = \angle\{\langle r_j, s(t-\tau)\rangle\}.$$

The updated received signal after detecting the first active sensor is

$$r_j^{(1)} = r_j - \hat{I}_{j,1}\, s(t - \hat{\tau}_1)e^{j\hat{\theta}_j},\ j = 1, ..., J.$$

In general, we denote the updated received signal after detecting k sensors by $r_j^{(k)}$, $k = 0, 1, 2, ...$, where $r_j^{(0)} = r_j$ is the original received signal. We continue with the detection process after the kth step as long as the following criterion is met:

$$\max_X \sum_{j=1}^{J} \tilde{I}_j\ |\langle r_j^{(k)}, s(t - \tau)\rangle| > T, \tag{2.17}$$

where T is a threshold, whose choice is discussed in Section 2.4.1. If the preceding criterion is satisfied, then the $(k + 1)$th sensor is localized as follows:

$$\hat{X}_{k+1} = (\hat{\tau}_{k+1}, \{\hat{I}_{j,k+1}\}) = \arg\max_X \sum_{j=1}^{J} \tilde{I}_j\ |\langle r_j^{(k)}, s(t - \tau)\rangle|\,,$$

Violation of the criterion (2.17) leads to termination of the algorithm. The threshold T is chosen (see Section 2.4.1) such that the probability of missing a true peak and the probability of false alarms or false peaks under a noise-limited setting meet user-defined tolerances. The algorithm obtained thus is far from optimal and suffers from error propagation, but the alternative of joint ML location estimation is computationally infeasible.

2.2.4 Standard SAR Algorithm versus ML Algorithm

In this section, we compare the performances of the ML and standard SAR algorithms, and attribute the poor performance of the latter to the random phase terms $\{e^{j\theta_{j,k}}\}$. For 5 active sensors at locations seen in Figure 2.3, Figures 2.4

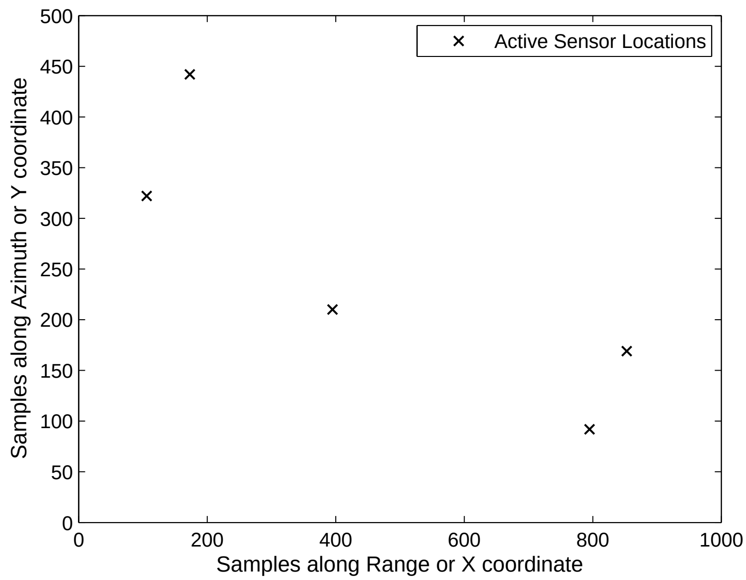


Figure 2.3: Active sensor locations for comparing standard SAR and ML localization. The sensor field is divided into a grid of 500 by 1000 samples in the azimuth and range dimensions respectively.

and 2.5 are the outputs of the processing using standard SAR and the modified SAR processing at an SNR = 4 dB respectively. Each image is the magnitude of the 2-D matched-filter output, $|\hat{\rho}(x, y)|$ in (2.9) and the decision statistic in (2.13), evaluated at the 500x1000 candidate locations in sensor field. The images are shown in grayscale with white representing the smallest and black the largest magnitudes. The elongated dark features in Figure 2.5 correspond to the responses of the active sensors to the beacon. In Figure 2.4, with the standard SAR processing, the responses of the active sensors are completely lost due to a combination of noise and destructive interference between snapshots that is caused by the random phase terms. However, in Figure 2.5, using the noncoherent techniques in Section 2.2.2, the active sensor responses are clearly distinguishable at their respective locations.

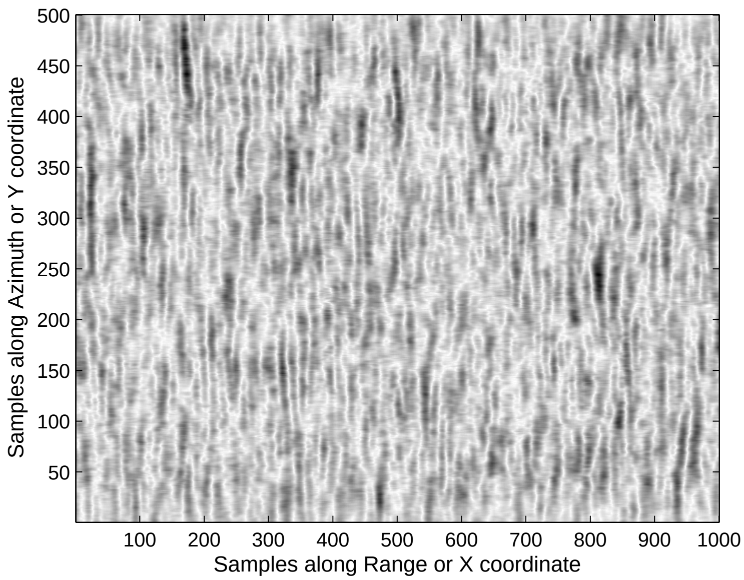


Figure 2.4: Output of standard SAR processing ($|\hat{\rho}(x, y)|$ from (2.9)) for scenario of Figure 2.3.

In addition to the preceding visual illustration of the limitations of standard SAR processing, we now compare sequential decision feedback using the standard and modified SAR decision statistics shown in Figures 2.4 and 2.5. We estimate the number of active sensors as a function of the threshold T, and plot it in Figure 2.6. With noncoherent modified SAR processing, the true number of sensors (5) are detected for any threshold between 50 and 60 (below $T = 50$, the estimated number of active sensors increases rapidly due to spurious noise peaks). On the other hand, with standard SAR processing, the difference between peaks due to signals and noise is difficult to detect, since the sum of signals with random

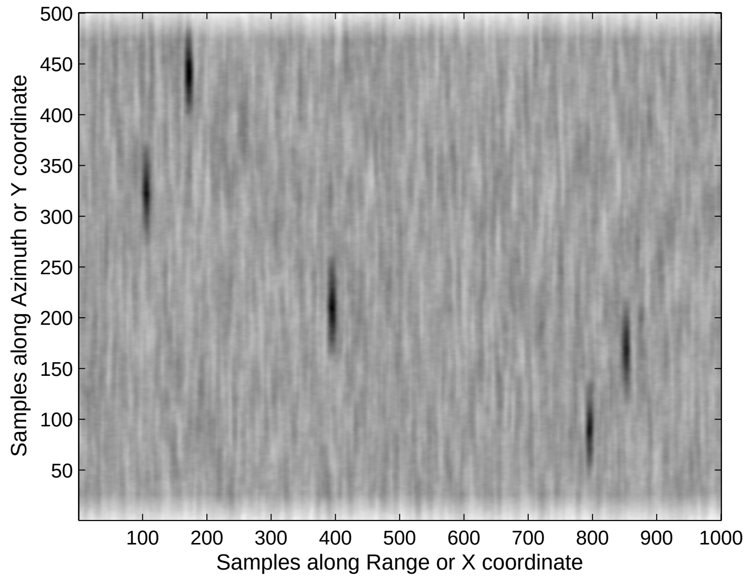


Figure 2.5: Output of ML algorithm ($|\hat{\rho}(x, y)|$ obtained by coordinate transformation of (2.13)) for scenario of Figure 2.3.

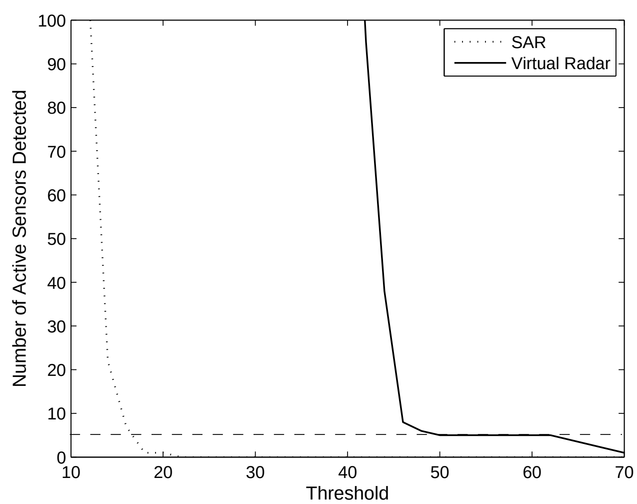


Figure 2.6: Comparison between standard SAR and ML techniques: Number of sensors detected is plotted versus the threshold. The dashed line shows the true number (5) of active sensors.

phases is noise-like. Thus, the estimated number of active sensors is dominated by spurious noise peaks for all values of T.

2.3 Noise-limited Localization Performance

In this section, we focus on noise-limited performance, in order to understand the dependence of the range and azimuth localization resolution on parameters such as SNR, antenna beamwidth, signal bandwidth and antenna beamshape. That is, we analyze the performance of the optimal ML localization algorithm (derived in Section 2.2.2) for a single sensor, ignoring inter-sensor interference.

2.3.1 Scale-Invariant System Dimensions

The description of system dimensions and characterization of system performance in terms of scale-invariant quantities enables prediction of the performance of another system with different physical dimensions, but the same relative dimensions. To this end, we introduce two normalizing parameters or so-called 'units' for the range and azimuth directions, respectively. For the range, the reciprocal of the nominal RMS signal bandwidth, $\mathcal{W}_{rms}$, expressed as a distance, $m_x = c/\mathcal{W}_{rms}$ where c is the speed of light, is used as the normalizing parameter. For the azimuth, $m_y = D$, the distance between successive snapshots, is used. The normalized distances along the 'x' and 'y' coordinates are $X = x/m_x$ and $Y = y/m_y$. The choice of m_x and m_y are based on the resolution analysis in

Sections 2.3.2 and 2.3.3. The simulation results in Section 2.3.5 are presented in terms of these dimensionless quantities.

2.3.2 Range Resolution

The analysis of the range resolution using the Cramer-Rao Lower Bound (CRLB) is only valid when the signal from an active sensor is acquired and the peak is within a half chip length of the actual location. The localization error depends on the distance between the true and estimated locations of the active sensors, and is therefore ill-defined, if the active sensor is not detected, or if there is a false detection. We present in this section a lower bound on the range resolution of the single sensor ML algorithm for an ideal brickwall AGF, under the assumption of sufficiently high SNR such that, the sensor is detected and its transmission acquired within one half-chip interval.

The received signal in (2.7) from a single active sensor is

$$r(x, y) = I(y - \overline{y})\, s(x - \overline{x})\, e^{j\theta} + n(x, y),$$

where n is AWGN, $\theta(y)$ is the random phase, I is the AGF, and $s(x) = s(ct/2)$ is the transmitted signal. We formulate range estimation as a 2 parameter estimation problem, where the range coordinate $\overline{x}$ and unknown random phase $\theta(y)$ are the parameters estimated. Since the sensor has been detected, $I(y - \overline{y})$ is, or the snapshots with a signal component are known. We use a notation where we drop the independent variables and write the received signal as

$$r = I.s.e^{j\theta} + n,$$

and

$$p(r|\overline{x}, \overline{y}, \theta(y)) = \frac{1}{2\pi\sigma^2} e^{-\frac{\|r - I\ s\ e^{j\theta}\|^2}{2\sigma^2}}.$$

The CRLB for the 2 parameter estimation, under the condition that both I and s are real, is

$$\begin{pmatrix} \frac{2\sigma^2}{\|I\|^2 \|s'(x)\|^2} & 0 \\ 0 & \frac{2\sigma^2}{\|I\|^2 \|s(x)\|^2} \end{pmatrix}.$$

Defining the RMS bandwidth of the signal $s(t)$, and noting that $s'(x) = \frac{2}{c}\ s'(t)$,

$$W_{rms}^2 = \frac{\int f^2 |S(f)|^2 df}{\int |S(f)|^2 df} = \frac{\|s'(t)\|^2}{4\pi^2 \|s(t)\|^2} = \frac{c^2 \|s'(x)\|^2}{16\pi^2\ \|s(x)\|^2},$$

where $|S(f)|^2$ is the power spectral density of $s(t)$, and c is the speed of light. Observing that $\|I\|^2 = N$ for the ideal AGF, which is the number of times the active sensor is illuminated by the beacon, and that $\frac{\|s\|^2}{2\sigma^2}$ is the transmit SNR per snapshot when $n \sim\ CN(0, 2\sigma^2)$, the CRLB for the error variance of the range estimate is

$$\sigma_{xx}^2 = \frac{c^2}{16\pi^2 W_{rms}^2\ \text{SNR}\ N}. \tag{2.18}$$

The right hand side of (2.18) has the dimension of c^2/W_{rms}^2 since the other factors are dimensionless. This motivates the choice of $m_x = c/\mathcal{W}_{rms}$ and $X = x/m_x$ in Section 2.3.1. The error variance of the dimensionless variable X is

$$\sigma_X^2 = \frac{\sigma_{xx}^2\ \mathcal{W}_{rms}^2}{c^2} = \frac{\mathcal{W}_{rms}^2}{16\pi^2 W_{rms}^2\ \text{SNR}\ N} \tag{2.19}$$

The use of delay estimation to determine the sensor range in the virtual radar system is evident from the dependence of the performance on the RMS signal bandwidth and transmit SNR in (2.19). The additional factor N accounts for the SNR improvement achieved by averaging out noise in the delay estimates from multiple snapshots, and N.SNR can be regarded as the effective range estimation SNR.

The CRLB provides insight into the tradeoffs between various system parameters by exposing performance trends as a function of these parameters. Since time-delay estimation is asymptotically efficient [71], the CRLB is achievable only at high SNR. However, we do not expect to attain the CRLB even at high SNR, since the CRLB does not account for the far-field approximation error (2.4): this is perhaps the most important effect at high SNR. Further, the localization error whose size is measured by the CRLB is not well-defined in the event of false alarms and misses, which become more likely at lower SNR.

2.3.3 Azimuth Resolution

We next evaluate an upper bound on the azimuth resolution for the ideal brick-wall AGF (the performance improves for more realistic Gaussian shaped beams, as shown in the numerical results later in this section). The azimuth estimate is quantized into bins whose size equals the distance D between snapshots, so that there is an irreducible quantization error taking values in $[-\frac{D}{2}, \frac{D}{2}]$. In addition, there are errors that can result from the choice of the wrong bin. In this analysis, we focus on characterizing the latter. For convenience, we label the correct bin as bin 0, as shown in Figure 2.7. We also assume that the range uncertainty has been resolved exactly, and that the active sensor has been detected (otherwise the localization error cannot be defined).

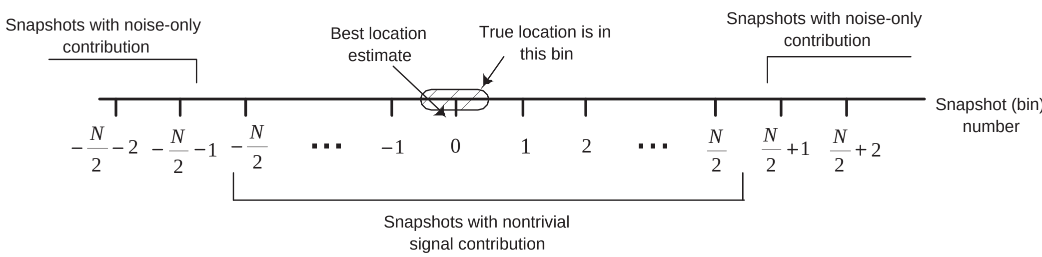


Figure 2.7: Random variables and errors in azimuth processing.

We define the *snapshot statistic* for snapshot l as

$$Z_l = |\int r_l(t)s(t-\tau)dt|, \tag{2.20}$$

where τ is the true round-trip delay and $r_l(t)$ is the received signal corresponding to snapshot l. Since an active sensor influences $N = \frac{2B}{D}$ snapshots, we see from Figure 2.7 that Z_l contains contributions from the desired signal (plus noise) for $|l| \leq \frac{N}{2}$, and contains contributions from noise only for $|l| > \frac{N}{2}$. The azimuth estimate is given by

$$\hat{l} = \text{argmax}_l \, Y_l,$$

where

$$Y_l = \sum_{k=l-\frac{N}{2}}^{l+\frac{N}{2}} Z_k, \tag{2.21}$$

is the accumulation of the snapshot statistics over a beamwidth centered around a hypothesized bin l (For simplicity, we make a slight change in notation, assuming that the normalized beamwidth, or the number of snapshots affected by a sensor, is $N + 1$, where N is even. The mean squared error thus derived is actually an upper bound for the case when the normalized beamwidth is N.) Since bin 0 is the correct bin, the azimuth error if bin l is chosen is lD. The mean squared azimuth localization error due to choosing the wrong bin can therefore be written as

$$E[|(\hat{l} - 0)D|^2] = E[|\hat{l}|^2]D^2.$$

We have

$$E[|\hat{l}|^2] = \sum_k k^2 p_k, \tag{2.22}$$

where

$$p_k = P[\text{choose bin k}] = P[k = \text{argmax}_l \ Y_l]$$

is the probability that bin k has the maximum accumulated statistic. Since the probability of attaining a maximum for an accumulated statistic containing noise-only snapshot statistics is very small, we limit the summation in (2.23) for $|k| \leq \frac{N}{2}$ (for $|k| > \frac{N}{2}$, Y_k has contributions from noise alone). Thus, we wish to evaluate

$$E[|\hat{l}|^2] = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}} k^2 p_k = 2 \sum_{k=1}^{\frac{N}{2}} k^2 p_k, \tag{2.23}$$

by symmetry. For $k \neq 0$, we can bound p_k by using pairwise comparison with the accumulated statistic corresponding to the correct bin 0 as follows:

$$\begin{aligned} p_k \;\leq\; P_k &= P[Y_k > Y_0] \\ &= P[Z_{\frac{N}{2}+1} + ... + Z_{\frac{N}{2}+k} > Z_{-\frac{N}{2}} + ... + Z_{k-1-\frac{N}{2}}]. \end{aligned}$$

Note that $Z_{\frac{N}{2}+1} + ... + Z_{\frac{N}{2}+k}$ is a sum of k snapshot statistics containing contributions from noise alone, while $Z_{-\frac{N}{2}} + ... + Z_{k-1-\frac{N}{2}}$ is a sum of k snapshot statistics containing signal as well as noise contributions. Under our model, the received signal in the lth snapshot is of the form

$$r_l(t) \;=\; s(t-\tau)e^{j\theta_l} + n_l(t)$$

signal present: sensor falls in beam.

In what follows, we label snapshot statistics corresponding to this "signal-present" scenario as $Z_l^{(s)}$.

$$r_l(t) \;=\; n_l(t)$$

noise only: sensor does not fall in beam.

We label snapshot statistics corresponding to this "noise-only" scenario as $Z_l^{(n)}$. Note that the WGN processes n_l are independent for different l, so that the snapshot statistics (conditioned on the sensor location) are independent random variables. Under these assumptions, it is easy to see from (2.20) that $Z_l^{(s)}$ are i.i.d. Rician random variables for "signal-present" snapshots, while $Z_l^{(n)}$ are i.i.d. Rayleigh random variables for "noise-only" snapshots.

Since the azimuth error scales with D, a convenient normalization factor along the 'y' direction is $m_y = D$. Defining the normalized azimuth coordinate $Y = \frac{y}{m_y}$, we infer from the preceding that (ignoring the bin quantization error)

$$\sigma_Y^2 \leq 2 \sum_{k=1}^{\frac{N}{2}} k^2 P_k, \tag{2.24}$$

where P_k, defined in (2.24), is rewritten below to emphasize the dependence on presence or absence of signal:

$$P_k = P[Z^{(n)}_{\frac{N}{2}+1} + ... + Z^{(n)}_{\frac{N}{2}+k} > Z^{(s)}_{-\frac{N}{2}} + ... + Z^{(s)}_{k-1-\frac{N}{2}}].$$

It remains to compute P_k. Computer simulations are a straightforward means of estimating P_k, since the values of k being considered are not very large. For moderately large k, an accurate alternative to computer simulations is the central limit theorem (CLT): since $Z_l^{(s)}$, $Z_l^{(n)}$ are independent random variables, the random variable $U_k = \left(Z^{(n)}_{\frac{N}{2}+1} + ... + Z^{(n)}_{\frac{N}{2}+k}\right) - \left(Z^{(s)}_{-\frac{N}{2}} + ... + Z^{(s)}_{k-1-\frac{N}{2}}\right)$ can be approximated as a Gaussian random variable with the same mean and variance. Define μ_s, σ_s^2 as the mean and variance of $Z_l^{(s)}$, respectively, and μ_n, σ_n^2 as the mean and variance of $Z_l^{(n)}$, respectively. (The dependence of these parameters on SNR has been suppressed from the notation.) Then

$$E[U_k] = k(\mu_n - \mu_s), \qquad Var(U_k) = k(\sigma_n^2 + \sigma_s^2),$$

and

$$\begin{aligned} P_k &= P[U_k > 0] \\ &\approx Q\left(\frac{\sqrt{k}|\mu_s - \mu_n|}{\sqrt{(\sigma_s^2 + \sigma_n^2)}}\right) \quad \text{CLT approximation.} \end{aligned}$$

In our computations, we use computer simulations for estimating P_k for $k \leq 5$, and the CLT approximation for $k > 5$. An alternative approach is to employ a Chernoff bound for P_k, but we find it to be less accurate than the method employed.

Since $\{P_k\}$ decay with SNR, so does the variance σ_Y^2. However, P_k also decays exponentially with k, as evident from the CLT approximation as well as from a Chernoff bound analysis. Thus, the first few terms dominate in the summation on the right-hand side of (2.24). Since the beamwidth of the antenna only affects the number of terms being summed, the azimuth localization error variance is expected to be insensitive to the beamwidth.

2.3.4 Running Example for Simulation Model

Our computer simulations are based on the following running example. The sensor field contains 2500 sensors randomly deployed on a 500 m x 500 m square grid. Active sensors that detect nearby events are randomly chosen from the deployed sensors while ensuring that there are no edge effects. The aircraft flies parallel to one side of the field at a distance of 1500 m and an altitude of 2500 m. When illuminated, each active sensor transmits a 11-chip Barker sequence, $s(t)$, using a BPSK constellation shaped with a square-root raised cosine pulse with 50% excess bandwidth. The carrier frequency of the sensor transmissions is 75 GHz. The side-looking antenna has a nominal physical beamwidth ($2B$) of 60 m (parabolic antenna diameter of 0.51 m), and the distance between snapshots,

$D = 1$ m (hence, $N = 60$). The antenna has either an ideal brickwall or Gaussian beam pattern as described in Section 2.1.2. The nominal root-mean-square (RMS) bandwidth of the transmitted signal is $\mathcal{W}_{rms} = 13$ Mhz and receiver sampling rate is 160 MS/s. The SNR at the receiver is defined as the ratio of the received power from each sensor to the noise power added at the collector per snapshot, i.e., $\|s\|^2/2\sigma^2$, when the noise is $CN(0, 2\sigma^2)$. Due to oversampling of the received signal, the noise power is appropriately scaled to maintain the correct SNR in the signal band. The nominal operating SNR is 2 dB using the threshold calculated in Section 2.4.1. The normalizing factors for the range and azimuth coordinates to achieve scale invariance are $c/\mathcal{W}_{rms} = 16.4$ m and $D = 1$ m respectively. The choice and motivation for these normalizing factors are discussed in Section 2.3.1.

2.3.5 Trade-offs between Parameters

In this section, ML localization performance in studied with emphasis on trade-offs between SNR, signal bandwidth, antenna beamwidth and antenna beamshape. A single active sensor is localized using the algorithm in Section 2.2.2 using the normalized RMS error (obtained from scale-invariant quantities) as the performance metric. However, the absolute RMS errors are also shown on an alternate y-axis in the plots. Due to the proportional relationship between antenna beamwidth and N (the number of times the sensor is illuminated), the two terms are used interchangeably, while the results are presented in terms of N. The nominal values defined in Section 2.3.4 were used for any parameter not explicitly mentioned in the simulation results.

Effect of Signal-to-noise Ratio

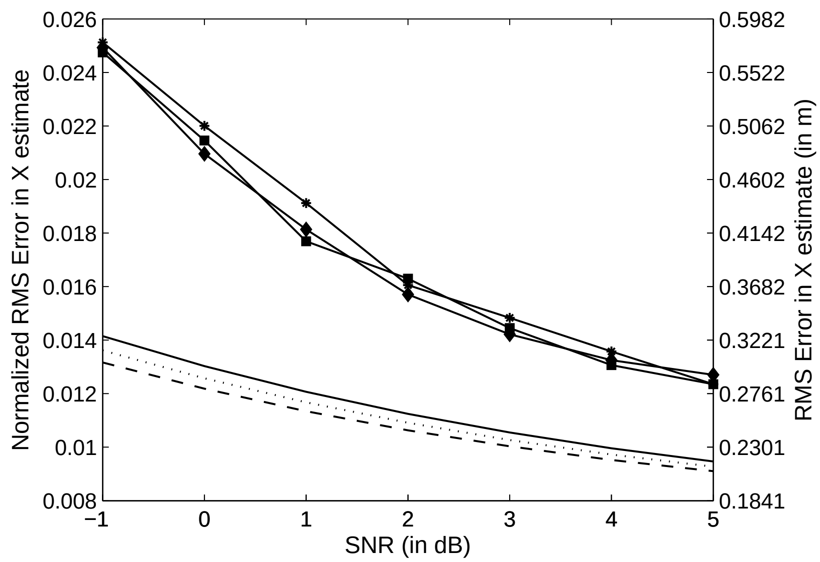


Figure 2.8: RMS estimation error in the range estimate versus SNR for different values of beamwidth: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

In Figures 2.8 and 2.9, the simulated RMS estimation error in the range and azimuth coordinates are plotted against SNR. The CRLB for the 'x' estimate and union bound (UB) for the 'y' estimate are also plotted for comparison. The bounds predict the trends in the error variance accurately, confirming the insights gained from the analysis in Sections 2.3.2 and 2.3.3. The RMS estimation error in both the range and azimuth coordinates decrease with SNR, and the analysis accurately

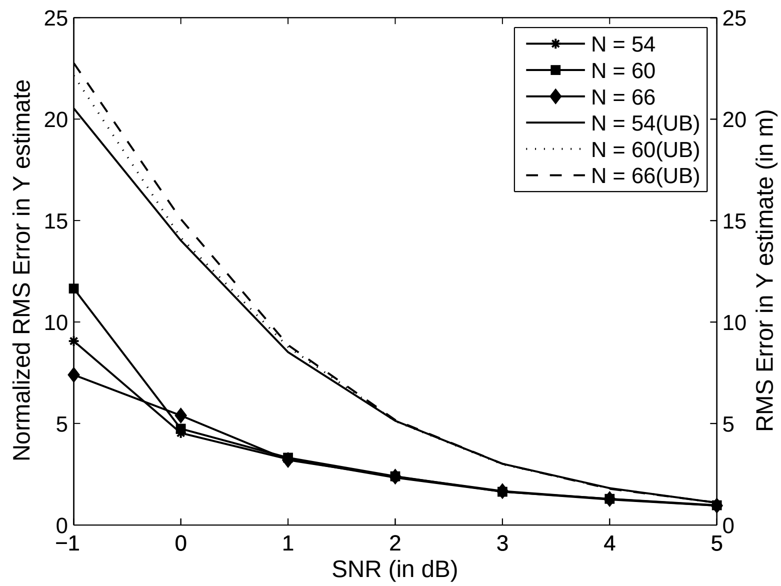


Figure 2.9: RMS estimation error in the azimuth estimate versus SNR for different values of beamwidth: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

predicts the expected performance improvement with SNR. Since delay estimation is asymptotically efficient, the RMS error is expected to achieve the CRLB at high SNR. However, we observe a gap to the CRLB in the simulations, which can be attributed to two main factors. First, the validity of the approximation error (2.4) becomes progressively worse as the beamwidth is increased. In the far-field approximation (2.3), the effect of the range on the azimuth coordinate of the sensor is neglected. At the nominal range $R = 3000$ m and beamwidth $2B = 60$ m, the worst case error due to the approximation is $B^2/2R = 0.148$ m, using (2.4) and recalling that the range estimate is the speed of light times half the round trip time. Second, our localization algorithm finds the range bin closest to the true sensor range at high SNR, which leads to residual quantization error. While this quantization error can be essentially eliminated by interpolation, we do not attempt to do this here. Instead, we note that the quantization and approximation errors do account for the gap to the CRLB for our system. For the system parameters in Section 2.3.4, the quantization interval is $\Delta = 0.578$ m, and leads to a quantization error variance of $\Delta^2/12 = 0.0278$ m^2. According to Figure 2.8, at 5 dB SNR, the observed mean squared error is 0.0808 m^2 and CRLB is 0.0455m^2. The gap to the CRLB is 0.0353m^2 which is well approximated by the sum of the quantization and approximation errors as $0.0278 + 0.0217 = 0.0495$m^2.

Effect of Signal Bandwidth

In Figures 2.10 and 2.11, the RMS estimation error in the range and azimuth coordinates is plotted versus the normalized RMS bandwidth of the transmitted signal for different values of N at SNR = 2 dB. The normalized bandwidth is the ratio of the true bandwidth and nominal bandwidth as defined in Section 2.3.2. The range estimate is inversely proportional to the RMS bandwidth of the transmitted signal. This is reflected in Figure 2.10, which shows the RMS error decreasing with the bandwidth and closely matches the CRLB in trend. On the other hand, the azimuth estimate is independent of the signal bandwidth and determined solely by the SNR. The agreement in trend between the estimation error in the 'y' coordinate and analytic results in Figure 2.11 also validates this analysis.

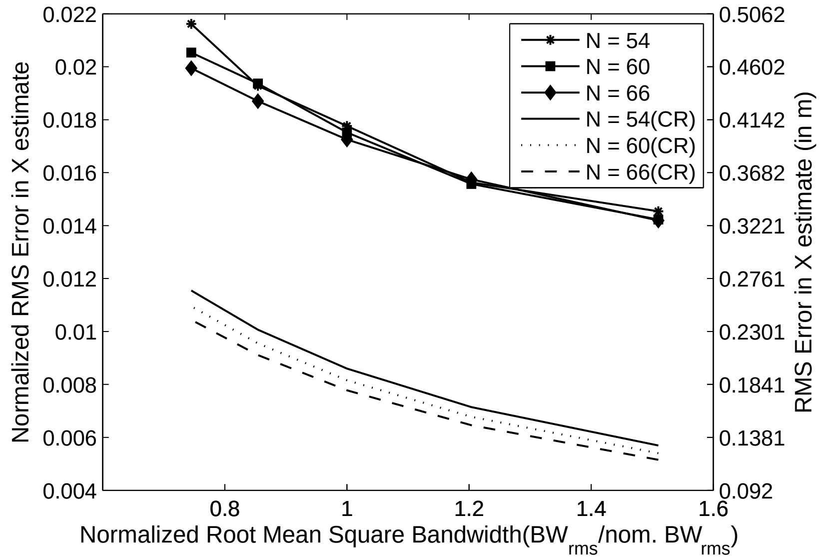


Figure 2.10: RMS estimation error in the range estimate versus normalized RMS bandwidth for different values of beamwidth: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

Effect of Antenna Beamwidth

According to the CRLB, the effective SNR for range estimation increases with N due to noise averaging over multiple snapshots, thereby improving performance. On the other hand for azimuth estimation, the error variances are dependent strongly on the SNR and weakly on N. The azimuth resolution is expected to be fairly insensitive to changes in antenna beamwidth. Figures 2.12 and 2.13, show the RMS estimation in the range and azimuth coordinates versus N for different values of SNR along with the analytical bounds, verifying that the dependence on the beamwidth is as expected.

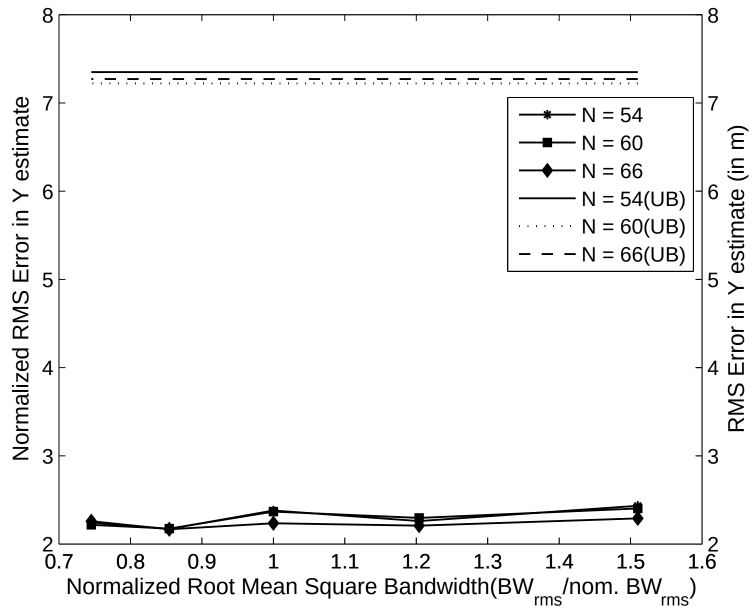


Figure 2.11: RMS estimation error in the azimuth estimate versus normalized RMS bandwidth for different values of beamwidth: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

Effect of Antenna Beam Shape

In Figures 2.14 and 2.15, the performance of a more realistic, smooth Gaussian beam, defined in Section 2.1.2, is compared against the idealized rectangular beam to gain insight into the effect of antenna beamshape. The Gaussian beams approximate the beamshape of parabolic antennae in the far-field fairly well. The beamwidth N of the rectangular and Gaussian beams are maintained equal, and the beamshapes are normalized so that the received power at the collector from each sensor is NP_s in order to isolate the effect of the beamshape. The Gaussian beam performs significantly better, and its better performance along the azimuth

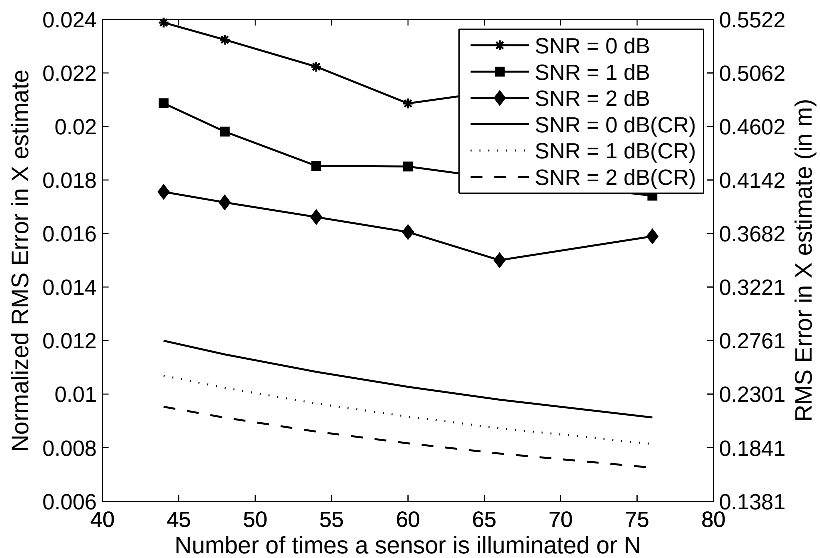


Figure 2.12: RMS estimation error in the range estimate versus beamwidth for different values of SNR: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

direction can be explained by the following two observations. First, the Gaussian beam has better autocorrelation properties (sharper autocorrelation peak) than the rectangular beam. Second, with the rectangular beam when the 'y' coordinate of a sensor lies between snapshots j and $j+1$, the received signal at the collector is exactly the same in the absence of noise, and this ambiguity causes performance degradation. On the contrary, with the Gaussian beam since the received power at various snapshots is a function of the sensor location, each location has a distinct received signal in the absence of noise, which leads to better performance. The improvement in the range estimate is mainly due to the improvement in the azimuth estimate. Although the algorithm is designed assuming that the range

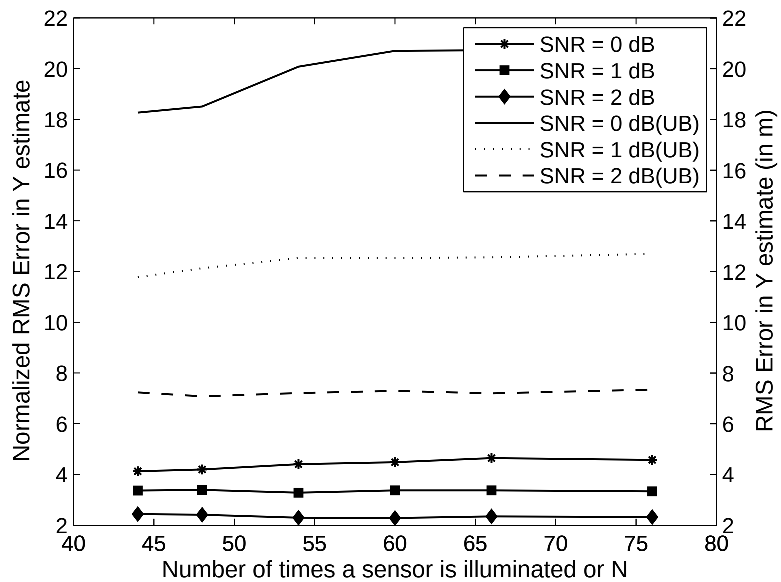


Figure 2.13: RMS estimation error in the azimuth estimate versus beamwidth for different values of SNR: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

and azimuth estimate are decoupled (due to (2.3)), in reality they are not. This result also justifies the study of the performance of the rectangular beamshape as a worst case scenario and an upperbound on performance in a practical setup.

2.4 Performance with inter-sensor interference

We now evaluate the performance of the decision-directed localization algorithm in Section 2.2.3. We first establish a termination criterion for the algorithm based on our analysis of noise-limited performance. The performance of the decision-directed algorithm and the optimal joint ML localization are compared

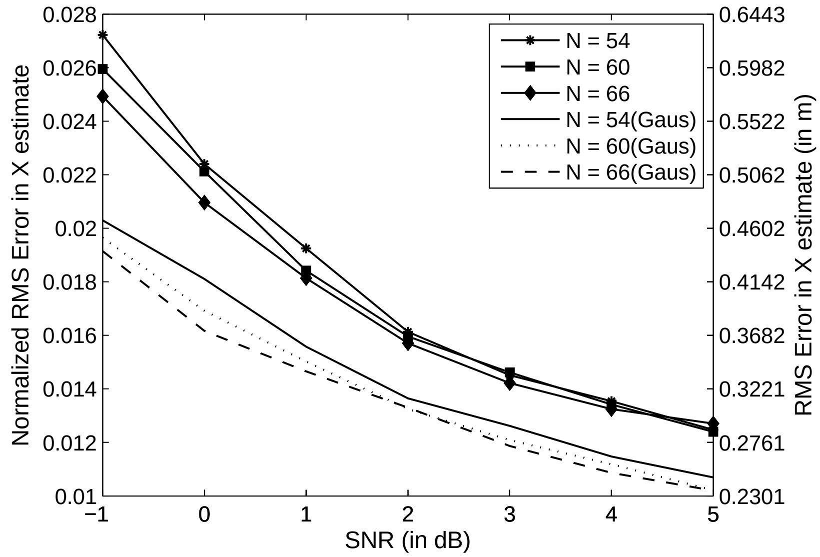


Figure 2.14: RMS estimation error in the range estimate versus SNR for different values of beamwidth for rectangular and Gaussian beams: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

for a small two-sensor example (the joint ML algorithm is too computationally complex for a larger number of sensors). There is a moderate penalty due to the loss of optimality, but this appears to be unavoidable, given the complexity of the jointly optimal algorithm. Finally, we evaluate the performance of the decision-directed algorithm for a dense sensor deployment.

For dense deployments, inter-sensor interference can cause significant degradation in detection performance. On the other hand, sensors that are close enough to interfere with each other may have correlated observations, and it may suffice to localize a subset of active sensors within such a cluster. This motivates us to define the concept of detection radius R_d: if the localization algorithm detects a

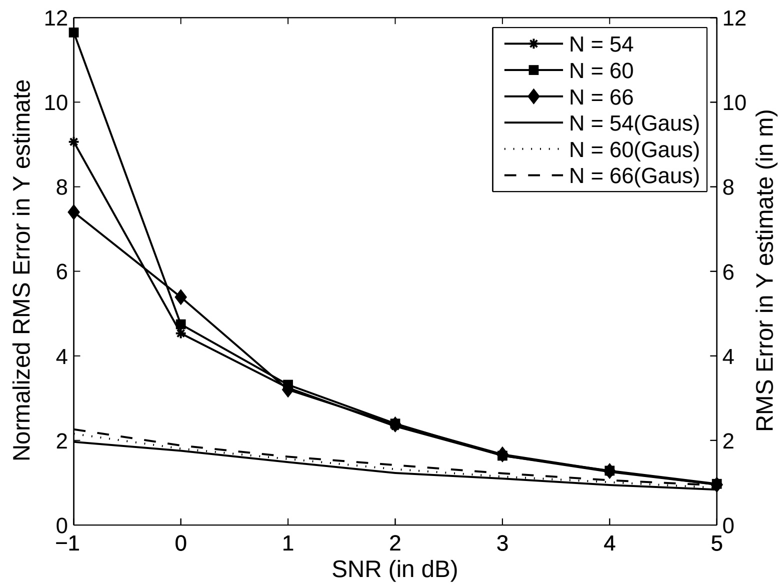


Figure 2.15: RMS estimation error in the azimuth estimate versus SNR for different values of beamwidth for rectangular and Gaussian beams: The RMS error in scale-invariant units and in meters are shown on the 2 y-axes.

sensor at a given location X, then it is deemed to have been successful in localizing any active sensor within a radius R_d of X. In practice, one might set R_d based on the anticipated spatial correlation in the sensor readings. Thus, a *miss* occurs if the decision statistics for *all* locations within radius R_d of an active sensor are below a threshold. A *false alarm* occurs at a specific location if its decision statistic exceeds the threshold, and it is not within R_d of an active sensor. In our numerical results, the sensors are deployed using a uniform distribution to achieve a density of 1 m^{-2} and R_d is measured in meters. In Section 2.4.1, where we determine the threshold based on a noise-limited analysis, we choose a small value of $R_d = 0.5$. In Section 2.4.3, where we investigate interference-limited performance for dense deployment, we consider the effect of increasing R_d on the probability of miss, in order to understand how correlations between sensor observations can ease the task of localization.

2.4.1 Termination Criterion for Decision-directed Algorithm

We now describe a method to choose a threshold to terminate the sequential detection algorithm in Section 2.2.3, when it is used to localize an unknown number of sensors. Under the detection algorithm, we repeatedly search for maxima in the decision statistic (2.17) until the magnitude of the maximum is below a threshold. It is difficult to analyze the effect of uncancelled interference on the performance of the decision-directed algorithm, hence we set the threshold based on noise-limited performance (i.e., it suffices to consider a single active sensor when determining the threshold). Letting p_{miss} denote the probability of miss, and p_{fa} denote the probability of a false alarm, the tradeoff between p_{miss} and p_{fa} is characterized by the receiver operating characteristic (ROC), which plots p_{miss} versus p_{fa} along a curve parameterized by the threshold. We use this curve to read off the threshold corresponding to desired levels of miss and false alarm probabilities.

Receiver Operating Characteristic(ROC): We use the notation defined in Section 2.3.3, and compare the accumulated statistics Y_l versus a threshold T. Suppose that there is an active sensor at bin 0, and suppose that bins $l \in L$ are within radius R_d of bin 0. A miss occurs if $Y_l < T$ for all $l \in L$. We obtain an upper bound on the probability of miss as follows:

$$p_{miss} = P(\max_{l \in L} Y_l < T) \leq P(Y_0 < T),$$

Recall that $Y_0 = \sum_{l=-N/2}^{N/2} Z_l^{(s)}$ (for a normalized beamwidth of $N+1$), where $Z_l^{(s)}$ are i.i.d Rician random variables, each with mean μ_s and variance σ_s^2. Thus, for moderately large N, Y_0 can be approximated as a Gaussian random variable with the mean $(N+1)\mu_s$ and variance $(N+1)\sigma_s^2$. Replacing $N+1$ by N for notational convenience, we obtain the following approximation for p_{miss}:

$$\begin{aligned} p_{miss} &\leq P(Y_0 < T) \\ &\approx Q\left(\frac{N\mu_s - T}{\sqrt{N}\sigma_s}\right) \quad \text{CLT approximation.} \end{aligned} \tag{2.25}$$

We now compute an approximation to p_{fa}, assuming that active sensors have been detected and cancelled. In this case, the accumulated decision statistic Y_k at a location k which is not within R_d of an active sensor is a sum of decision statistics due to "noise-only" snapshots: $Y_k = \sum_{l=k-N/2}^{k+N/2} Z_l^{(n)}$. Under a central limit theorem approximation, and replacing $N+1$ by N for notational convenience as before, we obtain that

$$p_{fa} \;=\; P(Y_k > T) \;\approx\; Q\left(\frac{T - N\mu_n}{\sqrt{N}\sigma_n}\right). \tag{2.26}$$

The parameters μ_s, μ_n, σ_s, and σ_n, which determine the ROC, depend on the operating SNR alone, apart from an arbitrary scale factor (if both signal and noise are scaled by a factor a, then all of the preceding parameters, as well as the threshold T, scale by a, but the SNR and ROC remains unchanged). Figure 2.16 shows the ROC for several values of SNR. The ROC is parametrized by the threshold T: increasing the threshold increases p_{miss} and decreases p_{fa}, and vice-versa when decreasing the threshold.

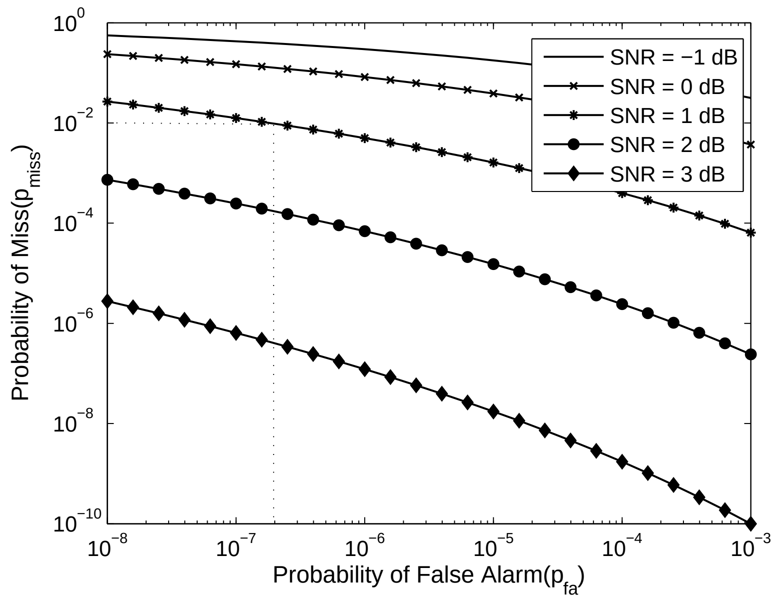


Figure 2.16: Receiver Operational Characteristic: p_{miss} vs p_{fa} for different values of SNR under the noise-limited setting.

The operating point, in terms of SNR and threshold, is determined by user-defined tolerances for miss and false alarm. While p_{fa} is the probability of false alarm at a given location, what is experienced by the user is the false alarm rate (FAR), defined as the probability that there is a false alarm at *some* location which is not within R_d of an active sensor. If there are M_{cand} candidate locations in our discrete grid, a union bound on the FAR is given by

$$\mathrm{FAR} \leq M_{cand}\, p_{fa}. \tag{2.27}$$

Example choice of threshold: Consider user-defined tolerances for the probability of miss and the false alarm rate as follows: $p_{miss} \leq 10^{-2}$ and $\mathrm{FAR} \leq 10^{-1}$. The sensor field is discretized into $M_{cand} = 500000$ candidate locations for running the localization algorithm, so that the requirement on FAR translates to $p_{fa} \leq \mathrm{FAR}/M_{cand} = 2 \times 10^{-7}$. From the ROC, the operating point must lie within the area enclosed by the dotted lines, and this provides the minimum operating SNR = 2 dB. To achieve the desired $p_{miss} = 10^{-2}$, at this operating SNR of 2 dB (since energy is at a premium, the lowest possible SNR is chosen), a threshold T is chosen using (2.25). This threshold and operating SNR are used to study performance of the suboptimal scheme in Section 2.4.3.

While we have used upper bounds on p_{miss} and FAR in the preceding formulation, this is still not sufficient to offset the effect of inter-sensor interference for a dense deployment. Thus, in practice, it would be necessary to add a link margin to the operating SNR determined by the ROC above.

2.4.2 Decision-directed algorithm versus Optimal ML Localization

The suboptimal decision-directed detection procedure in Section 2.2.3 for multiple sensor localization was adopted due to the computational complexity of the optimal ML algorithm. In Figure 2.17, the performance of this suboptimal algorithm is compared against the optimal algorithm for the simple instance of two active sensors in the field (when the optimal algorithm is still computationally tractable). The RMS estimation error in the azimuth coordinate is plotted against SNR for the two algorithms. When there is no overlap between the received signals corresponding to the two sensors in either the range or azimuth directions, then there is no inter-sensor interference. In this case, the suboptimal and optimal algorithms are identical and have the same performance. Furthermore in our examples, the bandwidth of the transmitted signal is large enough to provide adequate resolution in the range (or 'x') direction. Hence, to study the effect of inter-sensor interference, we focus on the scenarios where the two sensors have the same 'x' coordinate and are closely enough spaced in the 'y' direction to cause interference.

The simulations were performed on the system described in Section 2.3.4 with all parameters at their nominal values. In Figure 2.17, the RMS azimuth estimation error in scale invariant units is plotted versus SNR for the two algorithms. As expected the optimal ML algorithm performs much better and the disparity increases with SNR. However, at SNR = 4 dB the performance loss is about one unit, which is acceptable, considering that the computational complexity of the optimal algorithm is many orders of magnitude larger. As with a single sensor, the performance for both algorithms improves with SNR. In Figures 2.18 and 2.19, two instances of multiple sensor detection and localization using the suboptimal

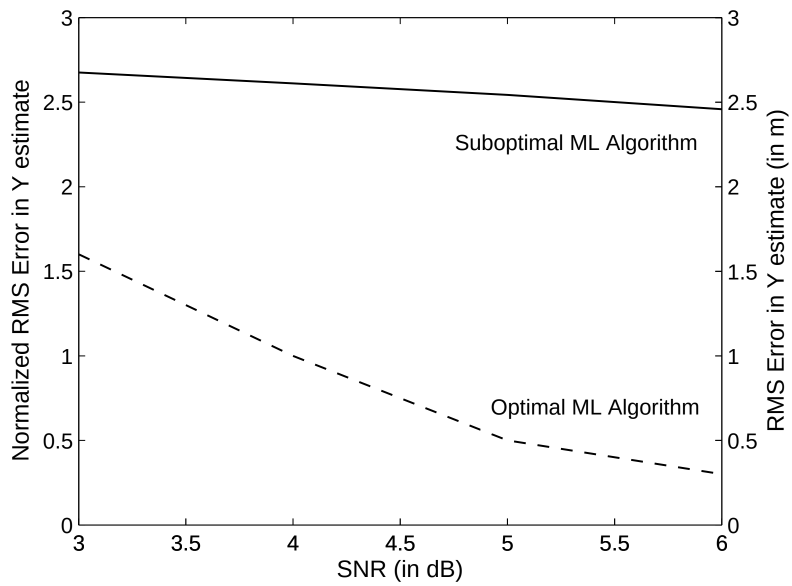


Figure 2.17: RMS estimation error in the azimuth coordinate versus SNR for the suboptimal and optimal ML algorithms with two active sensors.

algorithm are presented at SNR = 4 dB such that $p_{miss} = 10^{-2}$ and $p_{fa} = 2 \times 10^{-8}$ (false alarm rate of 10^{-2}).

2.4.3 Simulation Results for Dense Sensor Deployment

In this section, we investigate the algorithm performance in dense deployments, such as the scenario with 50 active sensors depicted in Figure 2.19. In such scenarios, we find through our simulations that a significant subset of active sensors are not detected, either due to destructive interference between the responses of nearby sensors (this is found to be the dominant effect in our simulations), or due to imperfect cancellation of the responses of the detected sensors. As mentioned

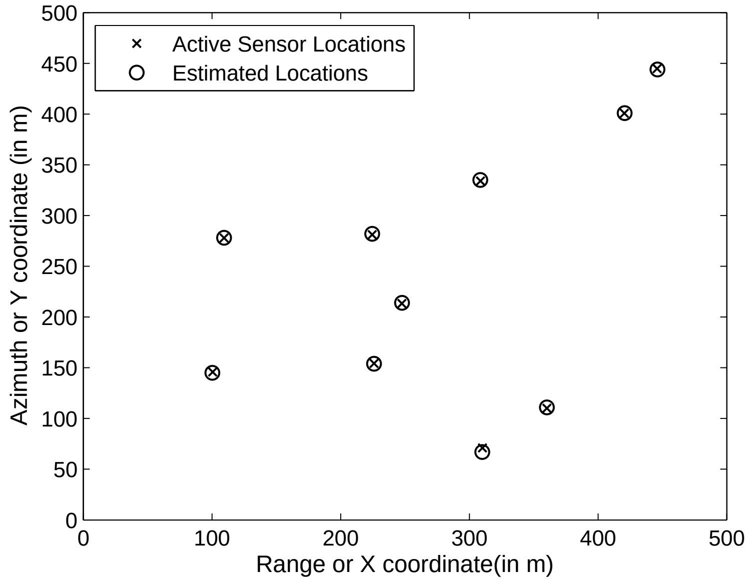


Figure 2.18: Detection and localization performance of suboptimal algorithm for 10 active sensors: The plot shows the 500m x 500m sensor field with true and estimated sensor locations.

earlier, it may be acceptable to detect a subset of the active sensors in a cluster if their observations are spatially correlated. We therefore explore the influence of the detection radius R_d on the probability of miss, as shown in the simulation results presented in Table 2.4.3. While we are interested in miss probabilities of the order of 1-10%, the threshold is chosen based on a noise-limited analysis for $p_{miss} = 10^{-3}$ in order to provision for the additional inter-sensor interference. From Table 2.4.3, we see that spatial correlation can significantly simplify localization, noting the significant reduction in p_{miss} as we increase R_d. For $R_d = 1, 2$ m, the observed p_{miss} increases with the number of sensors due to increase in the interference. However, for $R_d = 3$ m, the p_{miss} decreases marginally with the

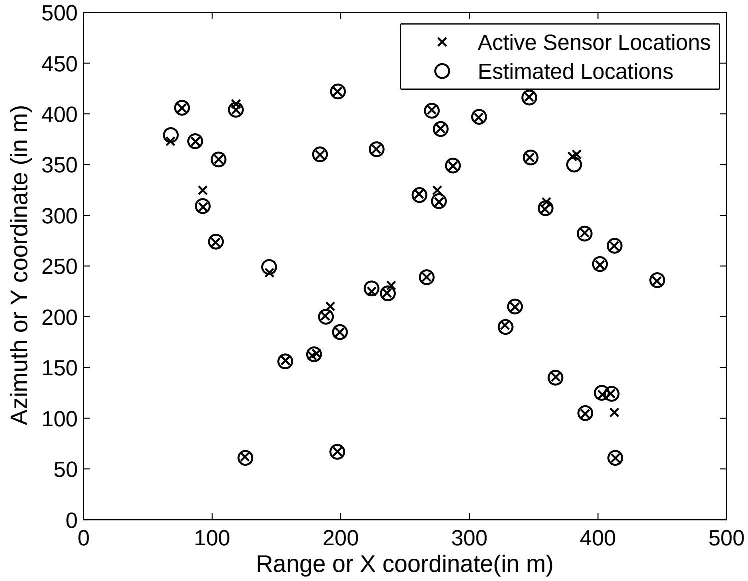


Figure 2.19: Detection and localization performance of suboptimal algorithm for 50 active sensors: The plot shows the 500m x 500m sensor field with true and estimated sensor locations.

number of sensors, since $R_d = 3$ m is large enough that some undetected non-interfering sensors randomly fall within R_d of a detected sensor,thus increasing our count of the number of 'detected' sensors. For instance, 2 sensors located about 3 m apart in the range direction do not interfere with each other. However, if one sensor went undetected due to noise alone, for $R_d = 3$ m, this sensor would be denoted as being detected, reducing the observed p_{miss}.

No. of Sensors	$R_d = 1$ m	$R_d = 2$ m	$R_d = 3$ m
10	0.0254	0.0270	0.0151
50	0.1168	0.0518	0.0143
100	0.1818	0.0719	0.0138

Table 2.1: The probability of miss decreases significantly with detection radius R_d.

2.5 Conclusions

We have shown, using an idealized model, that accurate localization is possible in large-scale imaging sensor nets with "dumb," severely energy-constrained, sensor nodes. The localization algorithm amounts to simple two-dimensional matched filtering for a single responding sensor, which can be extended to locate multiple sensors using successive interference cancelation. The localization algorithm was also shown to be applicable to a stationary collector with a mechanically or electronically steered antenna array and required only an appropriate choice of coordinate system to represent the received signal. Hence, the ML localization algorithm derived here is used as a significant block in the baseband algorithms in the imaging sensor net prototype presented in Chapter 3.

The localization accuracy in the range dimension can be improved by the increasing bandwidth of the beacon and in the azimuth dimension by decreasing the spacing between snapshots, when the effective synthesized aperture is sufficiently large. Moreover, most realistic antenna beam patterns give better azimuth performance than the ideal rectangular antenna beamshape. In dense deployments, where responses of proximal active sensors are lost to inter-sensor interference, we have shown that the decision-directed algorithm can locate at least one active sensor in each group of closely spaced active sensors. We also studied the detection of a responding sensor under this idealized one-off keying model in the presence of multiple sensor interference and characterized the ROC for the detector.

Chapter 3

Signal Processing for a Millimeter-wave Prototype

> *In theory, there is no difference between theory and practice. But, in practice, there is.*
>
> Jan L. A. van de Snepscheut (1953-1994)

In this chapter, we describe a millimeter-wave prototype of an imaging sensor net and report on the preliminary experimental results demonstrating the capability of this architecture to provide good localization resolutions while sustaining data rates of several kbps. The prototype consists of a stationary collector with a mechanically steered antenna and extremely rudimentary sensors, and is analogous to a system with an airborne collector as shown in Section 2.1.3. The initial implementation of the prototype was a "proof-of-concept" with the collector transceiver built using off-the-shelf components and a printed circuit board (PCB) sensor. Our goal is to gain a thorough understanding of the imaging sensor net system so as to design sensor and collector ICs, and develop the associated signal processing algorithms for data demodulation and sensor localization. In order to focus on our contribution to the prototyping effort, we present this work from a signal processing perspective and only describe the prototype hardware in sufficient detail to gain a functional understanding of its components. Moreover, we restrict our attention to the data demodulation and range resolution capability of the prototype since, the sensor can thereafter be localized in three dimensions by simple application of the localization algorithm presented in Chapter 2.

The main contributions in this chapter are:

1. We provide a system-level description of the millimeter-wave prototype hardware and discuss some of the design choices made.

2. Replacing the hardware components by functional blocks, we design algorithms for processing the collector receiver baseband outputs to demodulate the sensor data and locate the sensor in range.

3. The methodology, design and preliminary results of our experiments on this prototype to characterize the data demodulation and range localization performance are presented. These results indicate the feasibility of both the imaging sensor net concept and the chosen system implementation. However, the system performance is worse than predicted and reasons for this discrepancy are discussed.

The organization of this chapter is as follows. The various components of the millimeter-wave prototype, along with our system design choices, are described in Section 3.1. The collector receiver baseband algorithms are presented in Section

3.2. In Section 3.3, experimental results on the data demodulation and range localization performance are provided. Some of the primary sources of performance loss are discussed in Section 3.4 and finally, concluding remarks are made in Section 3.5.

3.1 The Millimeter-wave Prototype

We begin with an overview of the prototype hardware. While the abstractions from the localization model in Chapter 2 remain essentially applicable, significant modifications to the design are necessary to arrive at an implementation with truly simple sensor transceivers.

Instead of packetized reflection at the sensor, the collector transmits a periodic SS localization code that is continuously reflected by the illuminated sensor. The sensor also imposes low-rate data modulation on the electronically reflected collector beacon along with a frequency shift to distinguish the sensor reflection from backscatter. Correlation of the reflected beacon against the collector's copy of the location code creates a radar geometry for range estimation. Further, a stationary collector with a mechanically steered antenna sweeps the sensor field to produce a SAR-like three-dimensional response for each reflecting sensor that is utilized for sensor localization in azimuth and elevation.

The preceding paradigm is applicable to both *active* sensors, which amplify the beacon during reflection, and *semi-passive* sensors that do not. The term "passive" refers to the lack of amplification at the sensor, but is qualified by the term "semi" as there is, nevertheless, a power supply on the sensor to run the low-power baseband electronics. This terminology is the same as that in lower frequency RFID systems[16]. Accounting only for free-space propagation, the use of a *semi-passive* sensor incurs a $1/R^4$ loss, where R is the sensor range, whereas the loss in only $1/R^2$ for an *active* sensor. As shown in the link budget in Section 3.1.1, this enables a range of less than 500 m with the *semi-passive* sensor, while ranges of several kilometers are easily achievable with few milliwatts of transmit power at an *active* sensor. However, the design of the *active* sensor is more challenging, as it must effectively isolate the sensor transmit and receive chains to prevent oscillation. Hence, in this first generation prototype, we employ *semi-passive* sensors.

We chose to employ millimeter wave carrier frequencies in our prototype for the following reasons. A millimeter-wave design allowed us to miniaturize the sensor antenna, which remains the largest single component in the sensor even after integration of the sensor RF circuitry. There is also unlicensed spectrum available in the 60-70 GHz band that can be used for prototype development. Although this band suffers significant losses due to oxygen absorption, it is sufficient for the purpose of demonstrating the feasibility of a typical system. For long-range applications, we would avoid the oxygen absorption band, using, for example, millimeter wave frequencies above 70 GHz. Finally, millimeter-wave antennas with very narrow angular beamwidth can be obtained that are useful for spatially separating responses of multiple sensors.

After a discussion of the link budget in Section 3.1.1, the choice of the location code and sensor parameters are explained in Section 3.1.2. Finally, the collector transceiver and sensor designs are presented in Section 3.1.3 and Section 3.1.4 respectively.

3.1.1 Radio Link Characterization for Semi-passive and Active Sensors

The power budget for the communication link between the sensor and the collector is computed using the Friis' equation. For the *semi-passive* sensor, the transmitted beacon power is received at the collector after attenuation in the downlink and uplink, and the ratio of the received P_r and transmitted power P_t at the collector is

$$\begin{aligned} \frac{P_r}{P_t} &= \left(\frac{P_r}{P_t}\right)_{\text{up}} \left(\frac{P_r}{P_t}\right)_{\text{down}} \\ &= \left(D_{\text{TX}} D_{\text{sens}} \lambda_{\text{down}}^2 \, \frac{e^{-\alpha R}}{(4\pi R)^2}\right) \left(D_{\text{RX}} D_{\text{sens}} G_{\text{sens}} \lambda_{\text{up}}^2 \, \frac{e^{-\alpha R}}{(4\pi R)^2}\right), \end{aligned} \tag{3.1}$$

where D_{TX} and D_{RX} are the collector transmitter (TX) and receiver (RX) antenna gains, D_{sens} is the sensor antenna gain, G_{sens} is the sensor modulation loss, λ_{down} and λ_{up} are the downlink and uplink wavelengths, R is the distance between the sensor and the collector, and α is the atmospheric attenuation constant, which takes values in the range $6 - 16$ dB/km in the 60 GHz band depending on the atmospheric conditions. This computation does not account for losses such as mismatches, near-field effects and polarization. For a given data rate of r bits/s, the $E_b/N_0 = P_r/rFN_{\mathrm{th}}$, where N_{th} is the thermal noise density and F is the receiver noise figure.

On the other hand, when an *active* sensor is operated in the power-limited regime (fixed sensor output power), the uplink becomes the bottleneck from the perspective of power. Therefore, the power received at the collector is a fraction of the sensor output power P_s:

$$\frac{P_r}{P_s} = \left(\frac{P_r}{P_t}\right)_{\mathrm{up}} = D_{\mathrm{RX}} D_{\mathrm{sens}} \lambda_{\mathrm{up}}^2 \, \frac{e^{-\alpha R}}{(4\pi R)^2}. \tag{3.2}$$

It is seen from (3.1) and (3.2), that the *active* sensor has a more favorable $1/R^2$ fall-off in received power than the $1/R^4$ fall-off with a *semi-passive* sensor with sensor range R. With the collector transmitting at the maximum legal transmit power of 200 mW (determined by FCC regulation), the system can be operated at a range of 250 m with a *semi-passive* sensor sending data at 100 bps and E_b/N_0 of 11 dB. On the other hand, an *active* sensor transmitting 1 mW of power, can be employed up to a range of 5 km and communicate at 100 kbps with an E_b/N_0 of 11 dB, even if the oxygen absorption band is employed. These ranges are computed with the remaining parameters taking the following values: $D_{\mathrm{TX}} = 40$ dB, $D_{\mathrm{RX}} = 40$ dB, uplink frequency $f_{\mathrm{up}} = 60.5$ GHz, downlink frequency $f_{\mathrm{down}} = 60.55$ GHz, $G_{\mathrm{sens}} = -5$ dB,$\alpha = 10$ dB/km and receiver noise figure of 15 dB.

For the current experimental prototype setup, $D_{\mathrm{TX}} = 23$ dB, since a horn antenna is used at the collector transmitter (as seen in Figure 3.1), with all other parameters taking the values defined above. The maximum operating range for 100 kbps data rate at BER $= 10^{-4}$ is about 9 m using (3.1) with thermal noise.

3.1.2 System Parameter Choices

We discuss here the choice of the location code bandwidth, the sensor frequency offset and data rate of operation.

The range resolution of the system is dependent on the location code bandwidth as seen from the range estimation analysis in Section 2.3.2. Since a periodic location code is used, a sensor reflection causes periodic peaks in the output of the range correlator. Therefore, two sensors, whose propagation times to the collector differ by exactly one length of the location code, cause the same peak pattern and cannot be separated. This phenomenon is called "range ambiguity" and is a common shortcoming with radar-like systems that employ periodic signals for range estimation. In practice, prior information on permissible range , or other measurements such as received signal strength, may be used to overcome this problem.

In this experimental setup, a location code that is a pseudo-random noise (PN) sequence of length $N = 2^6 - 1 = 63$ chips (generated using a 6 bit shift register) with a bandwidth of 20 MHz (chip time $T_c = 50$ ns) is used. A difference in the arrival times of reflections of one chip is equivalent to a distance of $cT_c/2 = 7.5$ m, where $c = 3 \times 10^8$ m/s is the speed of light. Although a chip time corresponds to 7.5 m, range resolutions significantly smaller than the chip time can be achieved at sufficiently high SNRs, as illustrated in Section 2.3.5. The location code is repeated every $NT_c = 3.15$ μs and guarantees unambiguous operation in a sensor field with a dimension of 472.5 m, which is much smaller than our intended 9 m range of operation.

To maintain sufficient frequency separation between uplink and downlink (each has a bandwidth of 20 MHz), a sensor frequency shift of 50 MHz is chosen. Also, the sensor modulates differentially coded BPSK (DBPSK) data at 100 kbps on the beacon.

3.1.3 Prototype Collector

The prototype collector shown in Figure 3.1 consists of three modules: an up/down converter, steerable high-gain antennas, and algorithms for data recovery and localization. The collector signal processing is discussed extensively in Section 3.2 and hence, the discussion here focuses on the remaining components.

Up/Down Converter

The collector up/down converter translates the location code from baseband to the carrier frequency on the transmitter side and decomposes the received signal into the inphase (I-) and quadrature (Q-) channels at baseband on the receiver side. The location code is upconverted to the transmitter intermediate frequency

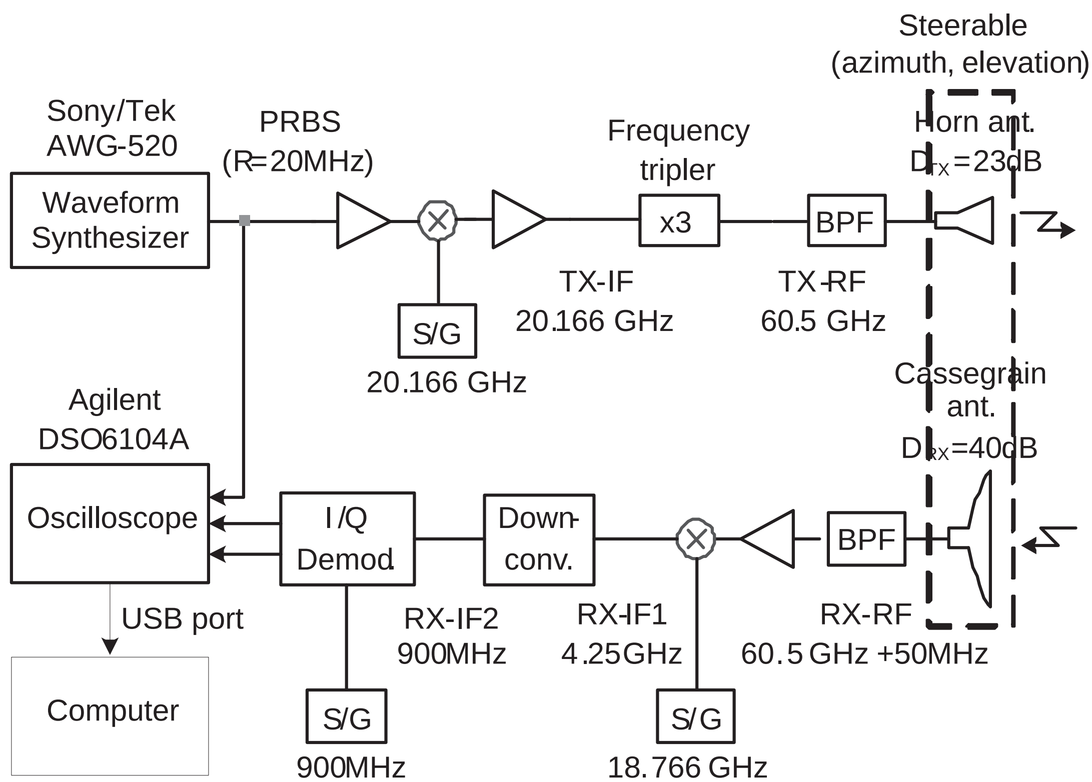


Figure 3.1: Block diagram of the 60 GHz prototype collector (Courtesy: Munkyo Seo)

(TX-IF) of 20.166 GHz (from a signal generator (S/G)) followed by a frequency triplication to the ultimate carrier frequency (TX-RF) of 60.5 GHz with a tripler output power of 7 dBm.

The reflection from the sensor at 60.5 GHz $\pm$ 50 MHz is received by a high gain collector antenna; the double sideband mixer used at the sensor produces a return at 50 MHz above and below the carrier frequency. A harmonic mixer and block-downconverter are used to translate the sensor return down to the first receiver

IF (RX-IF1 = 4.25 GHz) and subsequently the second receiver IF (RX-IF2 = 900 MHz), respectively. This corresponds to demodulating the lower sideband of the sensor return. Finally, an I/Q demodulator operating at RX-IF2 generates the I- and Q- channels of the baseband received signal that are digitized using a multi-channel oscilloscope for subsequent offline signal processing in MATLAB®. The transmitted location code is also captured in a third channel of the oscilloscope to provide a timing reference for range estimation (as discussed further in Section 3.2.1).

Steerable High Gain Antennas

The benefits of the high gain antennas used in this prototype are twofold. First, the high gain antennas have proportionally smaller beamwidths as evident from the relation

$$D \approx \frac{\pi}{\Theta_1 \; \Theta_2}, \tag{3.3}$$

where D is the antenna gain, and Θ_1 and Θ_2 are angular beamwidths (in radians) of the antenna along two orthogonal directions, respectively. The narrow angular beamwidths permit spatial separation of responses of multiple reflecting sensors. Second, the collector can transmit the beacon to sensors farther away with higher antenna gain at a given operating frequency. For a *semi-passive* sensor, with the $1/R^4$ power fall-off with sensor range, this additional gain is vital to prototype operation. The gain of the TX and RX antennas are $D_{\mathrm{TX}} = 23$ dB and $D_{\mathrm{RX}} = 40$ dB. The antenna patterns of the TX and RX antennas are circularly symmetric

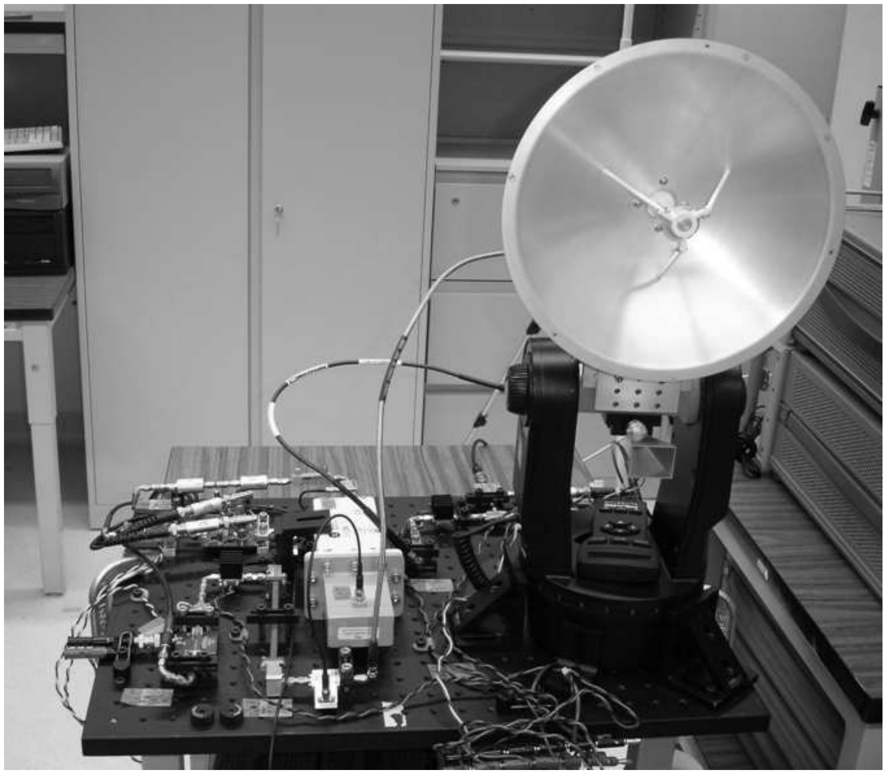


Figure 3.2: A photograph of the collector brassboard setup. (Courtesy: Munkyo Seo)

with half-power beamwidths of 14 degrees and 2 degrees, respectively. Notice, in Figure 3.1, that the TX chain has a horn antenna whereas the RX chain has a Cassegrain antenna. The TX and RX antennas are mounted on a motorized mount, using which the beacon can be steered in the azimuth and elevation with sub-degree accuracy. Figure 3.2 is a picture of the collector brassboard setup.

3.1.4 Prototype Sensor

At the sensor, the incident collector beacon is modulated with the local sensor data and frequency shifted (to distinguish it from direct ground returns at the collector), before being re-radiated back towards the collector. There is no attempt to recover the timing of the location code in the beacon to align the symbol transitions with location code chip transitions or to lock on to the carrier frequency to ensure phase synchronicity.

The block diagram of the 60 GHz prototype sensor is shown in Figure 3.3. The sensor consists of the RF circuitry implemented on low-cost Rogers 4005C substrate and a data generator built using digital logic ICs. A linearly-tapered open-slot antenna (LTSA) is chosen for its wideband operation and frequency-independent geometry that makes it robust to manufacturing errors. The designed antenna has > 10 GHz bandwidth centered at 60 GHz. This antenna provides an effective gain (accounting for losses) of 7 dB within a half-power beamwidth of 40 degrees.

The LTSA is terminated by a PIN diode. The impedance presented to the incident collector beacon is switched between two states with approximately 180 phase difference by turning the input bias on the diode on and off. Thus, the beacon is reflected with a relative phase shift of 0 or 180 degrees, which is equivalent to modulating the beacon with binary antipodal data (consisting of +1 or -1 symbols). Instead of implementing the data modulation and the frequency shift separately, the local data and 50 MHz digital clock for the frequency shift are

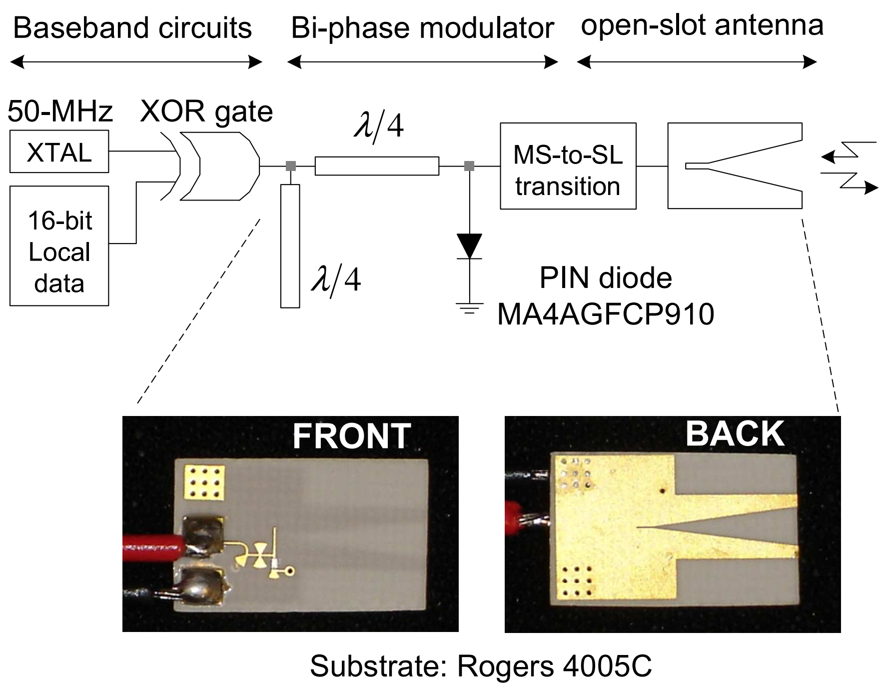


Figure 3.3: A prototype 60-GHz low-cost passive sensor with a wideband antenna and BPSK modulator. The size of RF circuitry is 15mm by 10mm. (Courtesy: Munkyo Seo)

jointly imposed on the diode bias via an XOR gate. The frequency shift is, hence, introduced digitally and results in a reflected signal at 50 MHz above and below the carrier frequency.

The digital logic on the sensor can be used to generate a repeating 16-bit sequence (selected by switches on the sensor) at the designated data rate (one of 1,10 or 100 kbps). In the current version of the sensor, the PIN diode draws a significant amount of current ($\sim$ 7 mA of dc current from a 3 V supply when forward biased), which retracts from our assumption of a low-power sensor. The dc power consumption will be significantly lower (in the μW), however, when the

RF circuitry is transitioned on to silicon with the exception of the LTSA, which must remain off-chip due to its considerable size.

3.2 Collector Baseband Processing

In this section, we develop algorithms for data collection and localization in the millimeter-wave prototype. The algorithm must operate on the I- and Q- channel data captured with the collector antenna pointing along different directions and extract the data modulated on the location code, in addition to locating the sensors in range, azimuth and elevation. Since the receiver outputs from multiple snapshots sampled at several times the chip rate need to be processed jointly at real-time speeds, non-decision directed "one-shot" estimation strategies are adopted whenever possible.

We formulate this coupled demodulation and location estimation problem accounting for two important factors. First, due to component tolerances, there is variation in the frequency offset applied at the sensor. This leads to residual frequency modulation when fixed IFs are used at the receiver. Second, the lack of synchronization between the symbol transitions of the sensor BPSK data and chip transitions of the location code makes the demodulation tightly coupled with the range correlation (despreading), which is unlike a standard SS system.

We present only the part of the baseband processing for data collection and range localization, but we indicate how the output of the range processing directly leads to the sensor location in azimuth and elevation through application of the algorithms in Chapter 2. Furthermore, we simplify the exposition by assuming a single reflecting sensor placed along the direction of maximum collector antenna gain.

A model for the signal received at the collector after reflection at the sensor is derived in Section 3.2.1. After initial delay and frequency estimation, the data is demodulated as shown in Section 3.2.2. Finally, the demodulated data and estimated frequency offset are used to estimate the location of the sensor in Section 3.2.3.

3.2.1 Received Signal Model

We now develop a collector received signal model for the purpose of designing the signal processing algorithm, while assuming ideal models for the hardware components and LOS propagation between the collector and sensor. The signals are represented in complex baseband notation using the second IF frequency (IF2) as reference, with time and phase origins referred to the collector. Thus, the real and imaginary parts of the complex baseband received signals are the in-phase (I-) and quadrature (Q-) channel outputs of the receiver. Since the output of the receiver is oversampled sufficiently, we present the algorithms assuming a continuous-time receiver output.

Let one period of the SS location code transmitted by the collector be $s(t)$ of length N chips and chip time T_c. The SS location code has a duration $T_{lc} = NT_c$

and each chip is encoded using BPSK. A sensor at a range R_S imposes a frequency shift f_{OFF} on the beacon and sensor DBPSK data is

$$b(t) = \sum_l d_l \; p(t - lT_b - \delta), \tag{3.4}$$

where $\{b_l\}$ are the BPSK symbols that are differential coded as $\{d_l = d^*_{l-1} b_l\}$, $p(t)$ is the symbol pulse shape, T_b is the symbol interval (duration of $p(t)$), and δ is timing offset between the collector and sensor clocks. The effective symbol pulse shape resulting from the *semi-passive* sensor design is an ideal brickwall of length T_b. Recall also that the sensor and the collector are not time synchronized. Suppose that the nominal receiver LO frequency f_{OFF} is Δf Hz from the actual frequency of the sensor reflection due to uncertainties in the magnitude of the sensor frequency shift.

Using the above notation and grouping together unknowns wherever possible, the signal at the output of the collector receiver chain can be written as

$$y_{\text{C-Rx}}(t) = A \underbrace{b(t)}_{\text{Data}} e^{\jmath 2\pi \Delta f t + \jmath\psi} \sum_k \underbrace{s(t - kT_{lc} - 2\tau)}_{\text{Range}} + n(t). \tag{3.5}$$

where A is the cumulative amplitude of the received signal, ψ is the phase offset between the collector and sensor LOs, τ is the one-way propagation time between the collector and the sensor, and $n(t)$ is the AWGN in the received signal from both the collector receiver and sensor transceiver. It is clear from (3.5) that the sensor data and sensor location information are tightly coupled in the received signal, and the residual frequency modulation must also be undone before the data or the range can be estimated. The angular location of the sensor is estimated by leveraging the variation in the received signal amplitude A with different orientations of the collector antennas. Although joint estimation of all the unknowns is optimal, it is computationally intensive.

The system parameters are chosen such that the bandwidth of the location code $\sim 1/T_c$ is much higher than than the residual frequency offset Δf and data rate $1/T_b$. Such a choice ensures that the phase change caused by the frequency offset over one period of the location code T_{lc} is not significant ($\Delta f T_{lc} \ll 1$), and that there are a large number of the location code repetitions in one bit period (i.e., $T_b \gg T_{lc}$). Since there are a large number of location code repetitions in one bit period, the phase coding in only a small fraction of location code repetitions is disrupted by symbol transitions (Recall that the chip transitions in the location code and bit transitions are not aligned).

The matched filtering with the location code can, thus, be performed without significant influence from the residual frequency offset and data demodulation. The received signal is first processed to estimate the residual frequency offset, and then the data. Soft estimates of the range can be obtained after removing the influence of the data and the residual frequency modulation. When data from multiple snapshots are available, these soft estimates can be combined across snapshots to estimate the ultimate sensor range and angular location. The block diagram of the complete signal processing algorithm is shown in Figure 3.4 and each block is elucidated below.

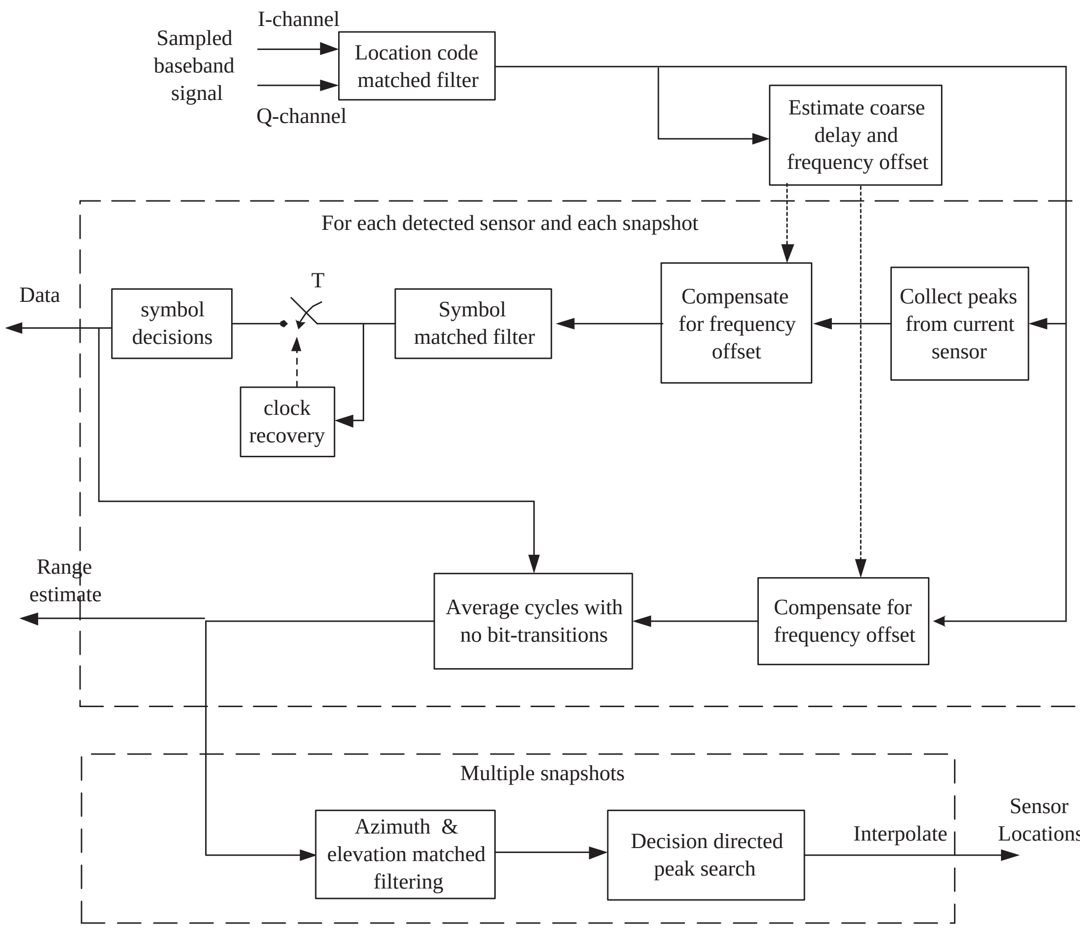


Figure 3.4: Block diagram of the Signal Processing Algorithm

Location Code Matched Filter: The received signal $y_{\text{C-Rx}}(t)$ is matched filtered against the location code $s(t)$ neglecting the effects of the residual frequency modulation and data:

$$\begin{aligned} y_{\text{MF}}(t) &= y_{\text{C-Rx}}(t) \otimes s(t) \\ &\approx A\, b(t) e^{\jmath 2\pi \Delta f t + \jmath\psi} \sum_k C_s(t - kT_{lc} - 2\tau) + n'(t), \end{aligned} \tag{3.6}$$

where $\otimes$ represents convolution and $C_s(t) = s(t) \otimes s(t)$ is the autocorrelation function of the location code. If a sensor reflection is present, a peak in the matched filter (MF) output corresponding to the sensor range (2τ) is obtained for each cycle of the location code and these peaks are exactly separated by T_{lc} seconds in the absence of noise, as seen from (3.6).

Location code Timing Reference: The definition of the propagation delay 2τ in (3.6) assumes that the location code is timed by an ideal clock with no drift and therefore, the periodicity of the transmitted beacon is exactly T_{lc}. To compensate for this possible drift, the propagation delay 2τ is measured differentially between the time each repetition of the location code is sent and the time it is received after reflection. To aid with this measurement, the transmitted location code is also captured on a separate oscilloscope channel $y_{ref}(t)$ (see Section 3.1.3), and is matched filtered against the location code. The peaks in this MF output provide the time when each copy of the location code was transmitted. To keep the notation simple, we retain the representation in (3.6) with each location code repetition beginning at $t = kT_{lc}$.

The peaks in the MF output $y_{\mathrm{MF}}(t)$ in (3.6) are scaled by complex factors that depend on the residual frequency offset and the modulating data during that location code cycle. Since, by design, there are multiple peaks in each symbol interval and residual frequency modulation period, this sequence of complex factors is an oversampled version of the data stream with the residual frequency modulation. When multiple sensors respond in a snapshot, each sequence of peaks in the MF output contains the data corresponding to different reflecting sensors and provides multiple access interference suppression as in a standard SS system.

Coarse Delay and Frequency Offset Estimation: In order to extract this sequence of complex factors, an estimate of the location of the correlation peaks is necessary. However, to reliably estimate the peak location and capture most of the peak energy, it is necessary to average noise out over multiple repetitions of the correlation peak, especially at low SNRs. But the presence of the unknown data and frequency offset do not permit this averaging. Since the data rate is low, the coarse delay and frequency can be jointly estimated. The MF output $y_{\text{MF}}(t)$ is coherently averaged for each of several discrete hypotheses of the delay and frequency offset, and the pair of hypotheses that captures most of the signal energy is chosen to correspond to the reflecting sensors. In essence, we generate a delay-frequency profile of the received signal and search for sensors in this two dimensional space:

$$\begin{aligned} Y_{\text{delay-freq}}(t,w) &= |\sum_k y_{\text{MF}}(t+kT_{lc})e^{-\jmath wk}|^2, \quad t \in [0, T_{lc}), \\ &\approx |A\sum_k \sum_l C_s(t+(k-l)T_{lc}-2\tau) \\ &\qquad e^{\jmath(2\pi\Delta f T_{lc}-w)k+\jmath\varphi} + n''(t)|^2, \; t \in [0, T_{lc}), \end{aligned} \tag{3.7}$$

where $\varphi = \psi + 2\pi\Delta f t$ and $n''(t) = n(t+kT_{lc})$. For a single reflecting sensor, the best pair of hypotheses for the coarse delay and and residual frequency offset is

$$(\hat{T}, \widehat{\Delta w}) = \arg\max \quad Y_{\text{delay-freq}}(t,w). \tag{3.8}$$

and the corresponding estimate of the phase offset is

$$\hat{\varphi} = \angle \sum_k y_{\text{MF}}(\hat{T} + kT_{lc}) e^{-\jmath \widehat{\Delta w} k}.$$

We can now compute the peak location and parameters of the residual frequency modulation as $\tilde{\tau} = \hat{T}/2$, and

$$\widehat{\Delta f} = \frac{\widehat{\Delta w}}{2\pi T_{lc}} \quad \text{and} \quad \hat{\psi} = \hat{\varphi} - 4\pi \widehat{\Delta f} \tilde{\tau}, \tag{3.9}$$

respectively.

Since the I- and Q-channel signals are already sampled in time, these samples naturally provide a set of coarse "delay bins" to search over. A periodogram is constructed for each delay bin, which is efficiently computed using the FFT. Moreover, as this FFT can be computed on the MF output spanning a length of time smaller than the bit interval, the effect of data modulation on the frequency and delay estimate can be kept to a minimum. This approach also provides an additional dimension to separate sensors over, for e.g., two sensors with similar ranges to the collector that would normally be hard to separate, could potentially be separated now, if they are sufficiently apart in frequency.

Collection of Peaks in the MF output: The coarse estimate $2\tilde{\tau}$ can now be used to extract the sequence of peaks in the MF output:

$$\begin{aligned} y_{\text{peaks}}[k] &= y_{\text{MF}}(kT_{lc} + 2\tilde{\tau}) \\ &\approx A' \, b(kT_{lc} + 2\tilde{\tau}) \, e^{\jmath 2\pi \Delta f (kT_{lc} + 2\tilde{\tau}) + \jmath \psi} + N[k], \end{aligned} \tag{3.10}$$

where $N[k]$ is a discrete white noise sequence and A' is the signal amplitude including the MF gain. The noise is white as the output of the MF is sampled every T_{lc} secs maintaining the independence between samples.

Frequency Offset Compensation: The residual frequency modulation in $y_{\text{peaks}}[k]$ can be canceled by appropriately rotating the phase of the peak samples using the estimated residual frequency offset:

$$\begin{aligned} y_{\text{data}}[k] &= y_{\text{peaks}}[k]\ e^{-\jmath 2\pi \widehat{\Delta f} k T_{lc} - \jmath\hat{\varphi}} \\ &= A'\ B[k] + N''[k], \end{aligned} \tag{3.11}$$

where $b(kT_{lc} + 2\tilde{\tau}) = B[k]$ is the datastream $b(t)$ sampled at a rate $1/T_{lc}$. We are, thus, left with the problem of decoding DBPSK data in AWGN. Our design choices ensure that the sampling rate $1/T_{lc}$ is larger than twice the data rate $1/T_b$, and hence, $y_{\text{peaks}}[k]$ reliably represents the data.

Next, data is demodulated from $y_{\text{data}}[k]$ in (3.11) (outlined in Section 3.2.2). Finally, using the demodulated data and estimated residual frequency offset, an estimate of the sensor range is obtained (described in Section 3.2.3).

3.2.2 Data Extraction

We describe, in this section, the symbol timing recovery and data demodulation algorithms.

Symbol Matched Filter: The datastream in (3.11) that is embedded in AWGN can be recovered by matched filtering with the symbol pulse shape $p(t)$. Define the sampled symbol pulse shape $P[k]$ as

$$P[k] = p(kT_{lc}) \quad k = 1, \ldots, N_b,$$

and $N_b = \left\lfloor \frac{T_b}{T_{lc}} \right\rfloor$ as the number of samples per symbol interval. The output of the symbol MF (SMF) is

$$y_{\text{SMF}}[k] = P^*[-k] \otimes y_{\text{data}}[k]. \tag{3.12}$$

But, the optimal times for sampling the SMF output are not known. Therefore, the data symbol clock must also be recovered from the SMF output.

Symbol Clock Recovery: In [72], a detector for a timing recovery loop in a BPSK system is presented that requires only two samples of the SMF output per symbol, one of which is used to make symbol decisions. The averaged error signal from the detector is used to adjust the symbol clock that samples the SMF output. The error signal is generated from the SMF output sampled at rate $2/T_b$, $y[k]$:

$$u_t[k] = y^*[k-1](y[k] - y[k-2]). \tag{3.13}$$

The clock produces the samples for optimal symbol decisions when the average of this error signal $u_t[k]$ is driven to zero. This approach for clock recovery also does not require phase synchronization, which provides added robustness against errors in canceling the residual frequency modulation.

However, in the prototype, the output of the SMF $y_{\text{MF}}[k]$ is already "sampled" at a rate $1/T_{lc}$, which is, nevertheless, much higher than the required sampling rate of $2/T_b$. Implementation of the timing recovery loop[72] in a discrete system (with some fixed sampling rate) requires interpolation and resampling to compute the SMF output values at the optimal sampling instants (see [73] for details). This approach is beneficial, however, only when the system sampling rate is comparable to the symbol rate.

We take a simpler approach here, in which the continuous-time timing recovery loop in [72] is implemented directly using the oversampled SMF outputs. The necessary error signal at a rate of two samples per symbol is generated from $y_{\mathrm{MF}}[k]$:

$$u[k] = y_{\mathrm{SMF}}[k - \lfloor N_b/2 \rfloor](y_{\mathrm{SMF}}[k] - y_{\mathrm{SMF}}[k - N_b]). \tag{3.14}$$

The error signal is averaged (recall the loop filter operates on the error signal at two samples per symbol) using a moving average filter of length N_{loop} as

$$u_{\mathrm{avg}}[k] = \sum_{l=1}^{N_{\mathrm{loop}}} u[k - l \lfloor N_b/2 \rfloor].$$

This $u_{\mathrm{avg}}[k]$ is a clock signal of period N_b for sufficiently large N_{loop}. As in a standard clock recovery loop, the zero-crossing of the clock signal with a negative slope is a stable operating point. Instead of interpolating the signal $y_{\mathrm{MF}}[k]$ to obtain the MF output at the precise zero-crossing of $u_{\mathrm{avg}}[k]$, we pick the sample index k_0 closest to the negative going zero-crossing during each period of the clock signal $u_{\mathrm{avg}}[k]$.

Let the sequence of optimal sampling instants recovered by the loop be

$$I = \{k + \epsilon : u_{\mathrm{avg}}[k] > 0 \text{ and } u_{\mathrm{avg}}[k+1] < 0, \epsilon = \frac{1 + \mathrm{sgn}(|u_{\mathrm{avg}}[k]| - |u_{\mathrm{avg}}[k+1]|)}{2}\}.$$

I will be an arithmetic progression with increments of about N_b, if the clock recovery is functioning properly.

Symbol Decisions: The BPSK data coded differentially as DBPSK symbols in $B[k]$ can be decoded as

$$\hat{b}_l = \mathrm{Re}\{\, y^*_{\mathrm{MF}}[I[l-1]]\; y_{\mathrm{MF}}[I[l]]\,\}, \quad l = 2, \ldots, |I|. \tag{3.15}$$

The differential encoding makes the demodulation resilient to errors in the residual frequency cancelation in (3.11).

3.2.3 Locating the Sensor in Range

The residual frequency estimate and demodulated data are now used to estimate the sensor range.

Data and Frequency Offset Compensation: For the purpose of data demodulation, the coarse delay was estimated using the delay-frequency profile. Although this approach is sufficient to extract the sample sequence containing the data, for range estimation, sub-sample delay resolution becomes necessary. As a first step, effects of the data and residual frequency modulation on the MF output (3.6) are undone:

$$\begin{aligned} y_{\mathrm{comp}}(t) &= y_{\mathrm{MF}}(t)\; \hat{b}(t)\; e^{-j2\pi\widehat{\Delta f}t - j\hat{\psi}} \\ &\approx A\sum_k \, C_s(t - kT_{lc} - 2\tau) + n'(t), \end{aligned} \tag{3.16}$$

where $\widehat{\Delta f} = \widehat{\Delta\omega}/T_{lc}$ and $\hat{\psi} = \hat{\varphi} + 4\pi\widehat{\Delta f}\tilde{\tau}$ from (3.10). Although it appears from (3.16) that $y_{\mathrm{comp}}(t)$ is equivalent to the output of the MF in the absence of data and residual frequency offset, (3.16) is strictly an approximation. The symbol transitions in the data alter the phase coding on the location code destroying its autocorrelation properties. The output of the MF in such situations is unsuitable for the purpose of accurately estimating τ.

Soft Delay Estimation: One approach to overcoming this problem involves compensating for the data and residual frequency modulation prior to the MF and repeating the location code matched filtering step in (3.6) on this corrected data. This approach clearly incurs a huge computational penalty due to the *repeated* MF operation. In order to avoid this added computational burden, we use the compensated MF output in (3.16), but drop all the cycles of the location code that contain a symbol transition. Since there are sufficiently many location code cycles in every symbol interval and the residual frequency modulation causes very little phase variation in a cycle duration T_{lc}, the cycles with no symbol transitions are equivalent to MF outputs in the absence of data and residual frequency modulation. Since the modulating data sequence $\hat{b}(t)$ has already been estimated, the instants when symbol transitions occur are also known.

Let $\mathcal{J}$ denote the set of indices of the transmitted location code cycles that do not contain symbol transitions,

$$\mathcal{J} = \{k : \ \mathrm{sgn}(\hat{b}(t)) = \epsilon, \ \forall \ t \ \in [kT_{lc}, (k+1)T_{lc}), \text{ and } \epsilon = \pm 1\}.$$

It is now simple to average multiple periods of the MF output that are aligned in phase to obtain one period of the MF output that can be used for localization:

$$\begin{aligned} y_{\text{range}}(t) &= \sum_{k\in\mathcal{J}} y_{\text{comp}}(t+kT_{lc}) \quad t \in [0, T_{lc}) \\ &= A\sum_{k\in\mathcal{J}}\sum_{l\in\mathcal{J}} C_s(t+(k-l)T_{lc}-2\tau) + n'(t+kT_{lc}), \;\; t\in[0,T_{lc}) \\ &= A\,|\mathcal{J}|C_s(t-2R_S/c) + \tilde{n}(t), \end{aligned} \tag{3.17}$$

where the variance of the noise $\tilde{n}(t)$ is $2|\mathcal{J}|\sigma^2$ and we make explicit the dependence of the averaged MF output on the sensor range R_S. The effective SNR for estimation of τ, thus, improves by a factor $|\mathcal{J}|$. The peak in $y_{\text{range}}(t)$ gives the estimate of the round-trip delay $2\hat{\tau}$ and hence, the range estimate is given by $\hat{R}_S = c\hat{\tau}$, where c is the speed of light. With sampled receiver outputs, the resolution in this estimate is limited by the sampling interval, but the range estimate can be improved by *interpolating* the MF output using the autocorrelation function of the location code $C_s(t)$.

Observe that $y_{\text{range}}(t)$ is identical to the output of the range correlation in our idealized model in Section 2.2 with the exception of the frequency offset, which has been already compensated for in (3.16):

$$y_{\text{range}}(t) = \langle r(t), s(t)\rangle. \tag{3.18}$$

As the range correlation output is available in the form of $y_{\text{range}}(t)$, an azimuth and elevation matched filtering across multiple snapshots, as indicated in Section 2.1.3, can be used to locate the sensor in three dimensions.

3.3 Experiments on the Prototype

In this section, the preliminary results of single sensor experiments performed on the prototype described in Section 3.1 using the algorithms in Section 3.2 are presented. The aim of these experiments were primarily to demonstrate the feasibility of the collector-driven imaging sensor net concept. In addition, the experiments were used to (a) verify the power-budgets for the communication link, (b) understand the bottlenecks involved in the development of the sensor ICs and (c) development and testing of signal processing algorithms. The experimental design and methodology chosen to characterize the performance of the prototype are described in Section 3.3.1. The results of the data demodulation and range estimation performance tests are presented in Section 3.3.2.

3.3.1 Experimental Methodology

We begin by describing the apparatus for capturing the received signal from the prototype. As seen in Figure 3.1, the I- and Q- channel outputs of the demodulator, and the transmitted location code (as a timing reference) are captured using the Agilent 6104 oscilloscope as an analog-to-digital converter (ADC). The received signals are sampled at 80 MS/s, which is four times the chosen location code chip rate of 20 MHz, at a resolution of 8 bits per sample. Since the oscilloscope does not operate in a real-time capture mode, the desired channels need to be sampled and stored in the oscilloscope memory first before the sampled signals can be transferred to a computer running MATLAB using a USB link. We are, therefore, limited in our capture capability by the oscilloscope memory depth to 50 ms of signal per channel at this sampling rate, and multiple captures of 50 ms became necessary to obtain sufficient number of bits for BER performance testing.

These preliminary experiments were designed to study the data collection and range localization performance of the prototype and the effect of one on the other. These single-snapshot experiments are important for understanding the working of the multi-snapshot azimuth and elevation processing. The experiments were performed with the sensor at different distances from the collector, which gives rise to links with different effective E_b/N_0.

We perform a variety of experiments at each sensor location. The system is operated in four settings with the collector either transmitting the location code on a carrier (upconversion 'on') or a single carrier frequency tone (upconversion 'off'), and with the sensor either modulating (data modulation 'on') or not modulating the beacon (data modulation 'off'). The frequency shift at the sensor, however, is always imposed on the reflection.

1. The Pilot Mode: There is no data modulation at the sensor and the location code upconversion is also turned off. This setup is used to measure the E_b/N_0 for the communication link at each range. The received signal consists of only the sensor frequency offset in AWGN. The parameters of the sinusoid, noise variance and finally the E_b/N_0 at the desired data rate can be estimated using the standard ML method in [74].

2. The Continuous-wave (CW) Mode: The data modulation at the sensor is turned on but the location code upconversion remains off. This setup is used to measure the BER at each range independent of the location code.

3. The Ranging Mode: This is complementary to the CW mode, and is equivalent to the idealized system in Chapter 2. The range estimate in this mode is used to observe the effect of data modulation on ranging.

4. The Nominal Mode: The system is operated in the standard setting with both the data modulation and location code.

3.3.2 Results

The experiments were performed at a data rate of 100 kbps. Since we focus on the data collection and range estimation performances, the collector was manually oriented to point directly at the sensor with the aid of a spectrum analyzer. We exploited the fact that the sensor reflection has maximum power when the collector antenna is pointed directly at the sensor. At each sensor location, 100000 bits of data were captured to measure BER with. The experiments were performed in a long rectangular hallway in the Engineering Sciences Building (ESB) that measures 2 m$\times$50 m. The sensor was located at distances between 3.6 m (12 ft) and 8.54 m (28 ft) from the collector in steps of 0.61m (2ft). The signal captured at the receiver also incurs circuit delays that adds an offset to the range estimates. This offset remains reasonably constant (there is some variation with temperature) and can be calibrated by estimating the range of a sensor at a known distance. The offset is calibrated using the range estimate at the highest SNR, and this computed offset is subtracted out of the range estimate.

In Figure 3.5, a comparison of the BER performance of the prototype under the nominal and CW mode at 100 kbps is shown. The E_b/N_0 for each sensor location is measured from the pilot mode experiments, with the smallest E_b/N_0 in the plot corresponding to the maximum range of 8.54 m. The measured BER for distances less than 5.5 m was less the 10^{-5} (no errors were found in 100000 bits). We observe that the BER measured in the nominal mode is only marginally worse ($<$ 1 dB) compared to the CW mode, and we can, therefore, conclude that data demodulation is not significantly affected by the presence of the location code. Nevertheless, the BER in the nominal mode is about 6-7 dB worse than the expected BER for DBPSK modulation[75] at the same E_b/N_0. We discuss the possible causes for this gap in Section 3.4.

In Figure 3.6, the E_b/N_0 measured using the pilot mode is plotted against the expected E_b/N_0 using the radio link characterization in Section 3.1.1 with thermal noise and receiver noise figure of 15 dB (tests of the receiver hardware showed that the noise figure is rather high). We, thus, verify that the collector received power does indeed fall as the fourth power of the range of the reflecting sensor. For instance, we can see, from Figure 3.6, that the E_b/N_0 drops by approximately 12 dB (16 times) when the range is doubled from 4 m to 8 m. Another factor that has been incorporated into the calculation is the power loss due to the "near-field"

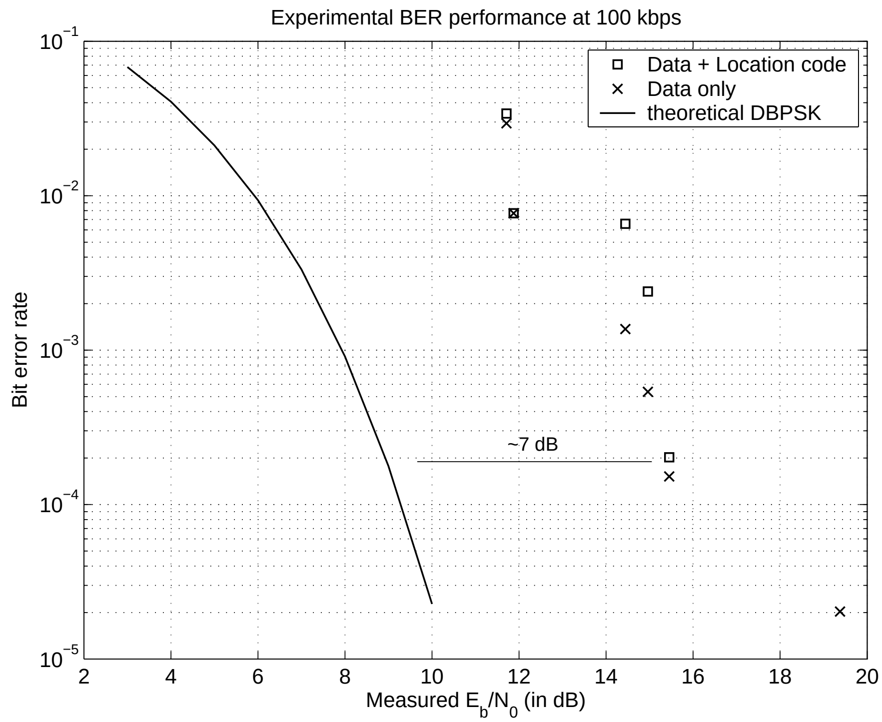


Figure 3.5: Prototype BER performance at 100 kbps: The observed BER under the nominal and CW modes were compared using the pilot mode to estimate the E_b/N_0.

effect. The Friis' formula for free space propagation holds good only in the far-field of the antenna. But, the operating sensors ranges in these experiments are not large enough to be considered far-field, and hence result in approximately 3 dB and 0.5 dB loss compared to the link-budget at 3.6 m and 8.5 m, respectively.

The mean estimated sensor range from the 20 captures is plotted against the actual sensor collector distance in Figure 3.7. The soft range estimate for each capture was obtained by averaging and interpolating 500 repetitions (free of bit transitions) of the location code using the algorithm described in Section 3.2.3.

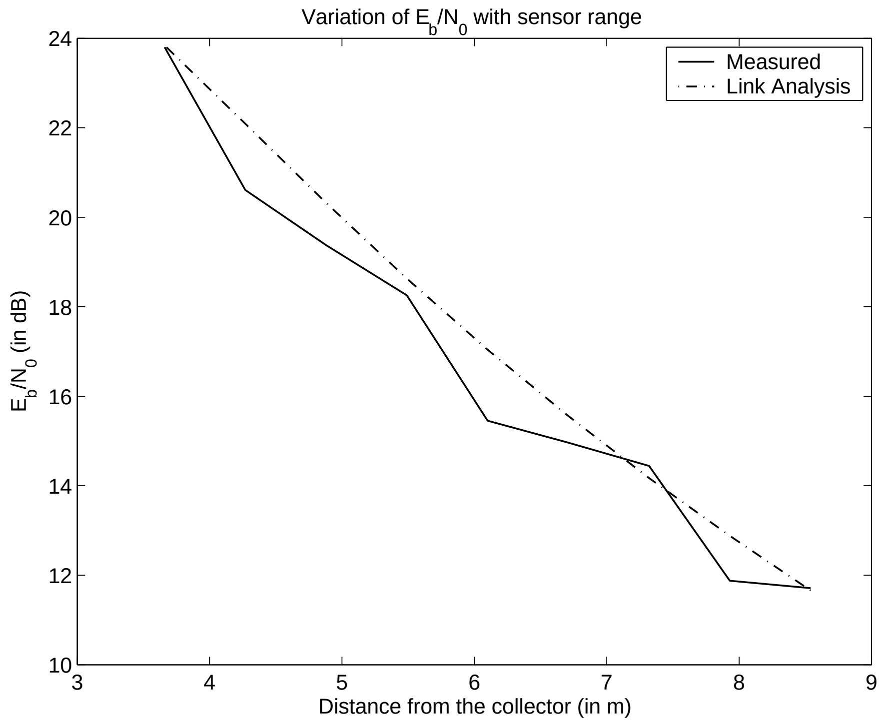


Figure 3.6: Verification of Radio Link Characterization: The E_b/N_0 measured from the pilot mode and computed using the link budget in Section 3.1.1 for different sensor ranges.

The offset due to circuit delays can be seen and remains roughly constant at different ranges as expected. Moreover, the mean estimated range from the ranging and nominal modes are in close correspondence, highlighting a negligible influence of data modulation on the range correlation and estimation.

The offset in the range estimate is computed to be 12.7 m using the mean range estimate at the highest E_b/N_0 (at 3.6 m). Using this offset correction, the RMS localization error for the ranging and nominal modes is shown in Figure 3.8. Our conclusion that the localization is not significantly affected by data modulation is

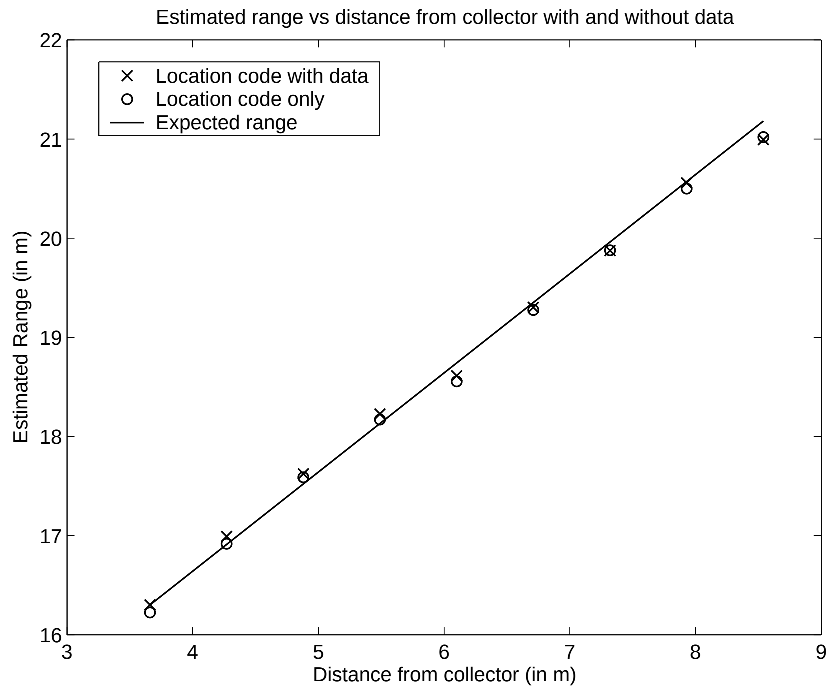


Figure 3.7: Variation of estimated range with sensor collector distance: Mean estimated sensor range versus true sensor collector distance under the ranging and nominal modes. The straight line (for reference) has unit slope and passes the through mean estimated range for 3.6 m range.

reinforced by this plot. We achieve about 10 cms range accuracy at all ranges with 300 interpolation points in every chip length. We observe that the performance is interpolation-limited (interpolation interval = 7.5/300 = 2.5 cms) and not noise-limited, which is the reason for the constant performance at both low and high SNRs.

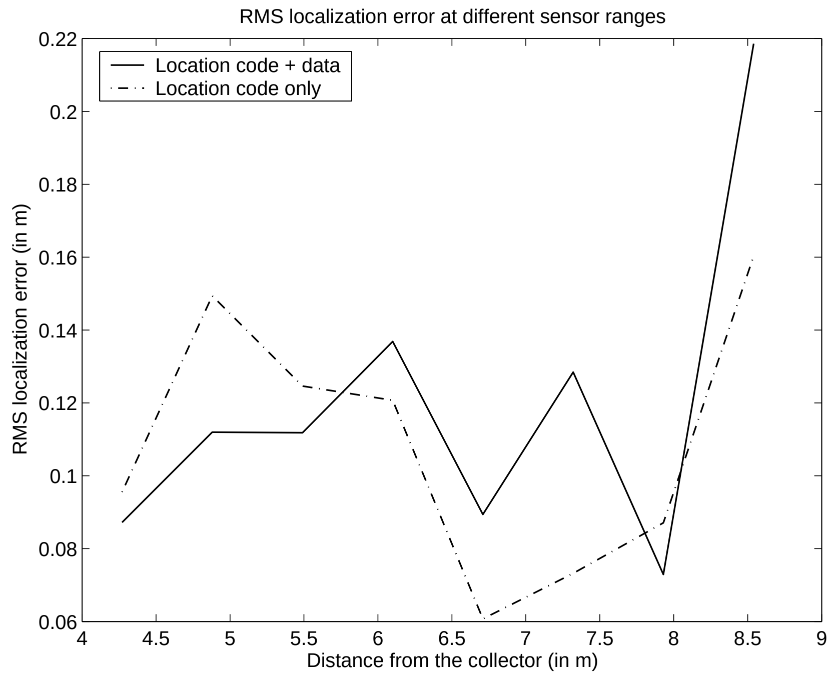


Figure 3.8: RMS localization error at different sensor locations: The RMS location errors in the ranging and nominal modes are plotted at different sensor locations. The offset in the range estimates is computed using the mean location estimate in the nominal mode for a range of 3.6 m.

3.4 Discussion of Issues and Possible Fixes

The effective loss of about 6-7 dB in the data demodulation in the experimental results can be attributed to a combination of several factors. We now discuss some of the issues observed in this system:

1. *Noise floor increase:* The frequency triplication of the 20 MHz location code in the collector upconverter shown in Figure 3.1 leads to spectral growth (intermodulation products) up to a bandwidth of 60 MHz. Since the uplink and downlink frequencies are separated by only 50 MHz (sensor frequency

shift) and the collector transmitter and receiver are co-located, there is leakage from the transmitter into receiver passband leading to an increase in the noise floor. The increase in noise floor was experimentally measured as the difference in the receiver output power between the ranging mode and the CW mode in the absence of a sensor. This experiment revealed that there was an increase of 2.2 dB in the noise floor due to leakage from the transmitter.

This issue can be easily addressed by increasing the frequency shift at the sensor or by replacing the tripler with a heterodyne upconverter. Although the current PCB sensor does not permit an increase in the frequency offset, the recently fabricated CMOS IC sensors are capable of producing frequency shifts of several 100 MHz. Future design of the collector IC will also eliminate the problem posed by the tripler.

2. *Interference due to direct return:* The direct return of the transmitted beacon from objects like walls and trees, has significantly more power than the sensor return, especially when the sensor is far away. The transmitted signal and therefore, the direct return has a 60 MHz bandwidth due to tripler. Further, the anti-aliasing filter at the input to the oscilloscope has a 3-dB bandwidth of about 25 MHz (Recall sampling rate is 80 MS/s). Hence, the frequency shift of 50 MHz is also not sufficient to prevent the direct return (which is 10s of dB above the signal level) from entering into the receiver passband and causing interference. The 2.2 dB increase in the noise floor

could, possibly, have a contribution from the direct return as well. This appears as spurious sensor peaks in the delay-frequency analysis disrupting the signal processing algorithm. However, in practice, the knowledge of the location of the aliased direct return can be used to eliminate these spurious peaks. As with the noise floor increase, an increase in the frequency shift should fix this problem.

3. *Distortion from frequency tripler:* The frequency triplication of the carrier with the location code also distorts the location code affecting its autocorrelation properties. Simulations using an ideal frequency tripler showed that correlation of the received signal with the undistorted location code template can lead to a loss of about 1 dB in the worst-case. The solution to the noise floor increase will alleviate this problem as well.

4. *Lack of symbol pulse shaping:* As described in Section 3.1.4, the sensor data modulation is performed digitally using an XOR gate. The effective pulse shape for this design is standard brickwall, which is not bandlimited. At a data rate of 100 kbps, the symbols have significant spectral components up to a bandwidth of about 300 kHz. The location code repetition rate (317 kHz) is the effective "rate" at which the symbol MF is sampled (see Section 3.2.2). Therefore, aliasing of the high frequency spectral components of the data symbols into the lower frequencies occurs, and causes inter-symbol interference. Unfortunately, this is the cost of having extremely minimalist

sensor design and the best immediate solution is to operate the system at data rates significantly lower than the location code repetition rate.

5. *Range Correlation loss:* In the signal processing algorithm, the received signal is correlated against the location code template neglecting the effect of the bit transitions. As mentioned earlier, symbol transitions destroy the correlation properties of the location code. At a data rate of 100 kbps, there are about 3.1 periods of the location code in every symbol interval. Since the correlation peaks of the location code are used as symbol samples for the bit demodulation, the loss of signal correlation peaks due to bit transitions can reduce the energy per bit especially at these high data rates. Moreover, since symbol matched filtering cannot be performed using 3.1 samples, about $10\log_{10}(3.1/3) = 0.2$ dB is lost by using a symbol width of 3 samples.

   Correlation of the received signal using subsequences (for instance, two halves) of the location code, and then combining these partial correlations in a decision directed manner (based on the observed bit transitions) can help recover a large fraction of the lost energy in the correlation peaks. There is, nevertheless, a higher computational and storage cost for manipulating MF outputs using multiple subsequences.

6. *Non-decision-directed signal processing:* The single shot estimation procedures adopted in the signal processing algorithm clearly perform worse compared to decision-directed or iterative procedures, although we are unaware

of the exact loss due to this suboptimality. A simple improvement to the algorithm would be to replace the timing recovery and symbol demodulation by the decision-directed scheme suggested in [75]. As suggested in the discussion on the range correlation loss, especially at data rates close to the location code repetition rate, an iterative loop involving range matched filtering-frequency estimation-data demodulation is necessary to operate close to the theoretical limits.

3.5 Conclusions

We have successfully demonstrated the feasibility of this “proof-of-concept” of an imaging sensor net with a stationary collector. Baseband processing algorithms for the collector receiver were developed for data demodulation and range localization. The output of the range localization can be used to locate the sensor in three dimensional space with the ML localization algorithm developed in the Chapter 2. We simplified the signal processing by choosing system parameters appropriately and achieved low algorithmic complexity by resorting to non-decision-directed schemes.

The proposed algorithms were used to perform indoor single sensor experiments on the prototype. The results, although promising, indicate a 6-7 dB loss in the BER performance. The range localization, on the the other hand, is limited by the interpolation resolution, since the localization SNR is high. The loss in performance can be attributed to a combination of several factor both in the hardware and in the signal processing, and possible solutions to several issues are discussed. The $1/R^4$ variation in the reflected power received from a sensor at range R and the radio link characterization were both verified experimentally. Nevertheless, these are but initial steps in our greater goal of deploying an imaging sensor net in the real-world and these preliminary results have provided us with many insights into designing the next generation of this prototype.

Chapter 4

Angle-of-Arrival-based Localization using a Network of Collectors

Our goal, in this chapter, is to gain fundamental insight into cooperative localization of a source by a network of geographically-dispersed receivers using AoA measurements. Such source localization problems have attracted a lot of attention particularly in the context of cellular systems, where the network of base stations (BSs) is required to locate a transmitting cell phone in emergency situations in accordance with the E-911 FCC mandate. The state-of-the-art localization solutions in cellular systems can be classified into the collector/BS-based or handset-based categories. In the former, the network of BSs locate the phone using time-difference-of-arrival (TDOA), AoA, time-of-flight (TOF), and multipath profiling while in the latter, TDOA is measured at the phone rather than the BSs. These techniques provide localization accuracies between 50 to 200 m depending on the environment.

In this chapter, we consider AoA measurements alone, since they only require that each receiver has a calibrated antenna array with known orientation and location. It is assumed that the receivers can coordinate to pool all the AoA estimates corresponding to a given source (e.g., using coarse timing synchronization between the receivers to associate AoAs from a given source based on time-stamps). These receivers can, therefore, be implemented at low cost without the need for the stringent time-synchronization as required by TOF and TDOA. Although RSS can also be measured using low-cost receivers, RSS measurements are known be notoriously unreliable.

Moreover, LOS between the receivers and the source is typically a prerequisite for accurate source localization. In practice, due to scattering and reflection at surfaces, NLOS multipath can cause large deviations (*outliers*) in the measurements that can result in poor localization performance, if not accounted for. Hence, we develop models and algorithms for the AoA measurements for handling NLOS scenarios considered here. Performance improvements possible by using additional forms of location information such as TOF or RSS measurements requires further exploration and is left as a topic of future work.

Although our initial investigation of this problem was in the context of a *sensor-driven* network, this work has applications in a wide variety of source localization applications such as sonar and multimedia systems. Therefore, we use the terms 'source' and 'sensor', and 'collector' and 'receiver' interchangeably. We also omit here a discussion of the receiver front-end algorithms for detection and acquisition of the source signal (see our prior work[76] for discussion), and instead focus on the AoA estimates generated by the receivers' array processing algorithms. We present a class of models representative of the AoA estimates obtained under different LOS and NLOS settings with typical estimation techniques, and use these models to develop localization algorithms.

Our main results are:

1. A sequential algorithm for aggregation of AoA estimates to estimate the sensor location under LOS scenarios (minimal multipath scattering) is presented. The collectors only need to exchange sensor location and covariance estimates, and the algorithm has linear complexity in the number of collectors. The location uncertainity grows with the size of the coverage area and the variance of the AoA estimates, but is reduced by increasing the number of collectors.

2. Since ML localization of the sensor under NLOS propagation requires exhaustive search of exponential complexity over all possible subsets of collectors, we propose a randomized algorithm for *outliers* suppression using the sequential algorithm as a building block that has complexity $O(MN)$, where M is the number of randomizations and N is the number of receivers.

3. The proposed algorithms are numerically shown to achieve the ML performance under different propagation models. We also illustrate with specific examples that a wideband system, with its capability to resolve the LOS and

NLOS multipath, along with the proposed outlier suppression algorithm, greatly outperforms a narrowband system, which is without this capacity.

This chapter is organized as follows. The AoA measurement models under different propagation environments are described in Section 4.1. The sequential algorithm for localization under LOS scenarios along with performance benchmarks are presented in Section 4.2. We derive an extension to the sequential localization to suppress NLOS AoA estimates in Section 4.3. The proposed algorithms are numerically investigated in Section 4.4. Finally, concluding remarks are made in Section 4.5.

4.1 Models of AoA Estimates

The collectors are equipped with an uniform linear array (ULA) to estimate the direction of arrival of the sensor transmission. In this section, we develop models to capture the effects of LOS and NLOS propagation environments on the AoA estimates produced by the collectors, and use these models to design network-level localization algorithms in Sections 4.2 and 4.3.

The classical AoA estimation techniques like MUSIC and ESPRIT[77] are based on LOS propagation from the sensor to the collectors, where the sensor is modeled as a point source. This point source propagation model is often unrealistic since they do not account for multipath scattering. A number of models for sources with local scattering and methods of AoA estimation under those models have been proposed [78, 79, 80, 81, 82], in which the propagation is characterized by a mean arrival angle (corresponding to the true bearing) and a spatial spreading parameter (quantifying the spatial extent of the multipath effects). The resulting AoA at the antenna array have been represented by a Gaussian [81] and Laplacian distributions [83]. On the basis of these ideas, we propose the following models:

LOS Model: We characterize the spatial spreading in LOS propagation scenarios by zero-mean symmetric finite variance "noise" models like the Gaussian and Laplacian. The AoA estimation under LOS propagation in the presence of additive white Gaussian noise results in zero-mean Gaussian errors, whose variance grows with the signal-to-noise ratio [84]. Additional degradation to the AoA measurements, is effected by local scattering in the vicinity of the sensor that, as mentioned earlier, also leads to zero-mean symmetric distributions. Therefore, the Gaussian and Laplacian error models can seamlessly transition between the scenarios with and without local scattering by varying the spatial spreading parameter. These models represent situations where the received signal has a strong LOS component along with limited scattering. For a sensor at location $\underline{X}$ along a true bearing $\theta(\underline{X})$, the Gaussian LOS model with local scattering is

$$p_{\text{Gaus}}(\hat{\theta}/\underline{X}) = \frac{1}{\sqrt{2\pi}\sigma(1 - 2Q(\frac{\pi}{2\sigma}))} \exp\left(-\frac{(\hat{\theta} - \theta(\underline{X}))^2}{2\sigma^2}\right) \quad \hat{\theta} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad (4.1)$$

where σ^2 is the error variance (spatial extent of scattering), $Q(t) = \int_t^\infty \exp(-t^2/2)$ $dt/\sqrt{2\pi}$ is the normal tail distribution and all angles are measured with respect to the antenna broadside. The Gaussian density is truncated as an antenna array can only differentiate between AoA within $\pm\pi/2$ of the direction in which the antenna broadside is pointing. Note also that as the variance σ^2 increases, the density become progressively long tailed and tends toward a uniform density. Similarly, the Laplacian LOS model with spatial spreading factor σ is a model with heavier tails than the Gaussian model:

$$p_{\text{Lap}}(\hat{\theta}/\underline{X}) = \frac{1}{2\sqrt{2}\sigma \exp(-|\pi/(\sqrt{2}\sigma)|)} \exp\left(-\frac{\sqrt{2}|\hat{\theta} - \theta(\underline{X})|}{\sigma}\right) \quad \hat{\theta} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \tag{4.2}$$

Worst-case NLOS model: In an extreme case of a NLOS propagation environment, the LOS path to a collector might be blocked, for instance by structures like walls or trees, and as a result, the received signal is composed exclusively of multipath components that are "far" from the LOS path. AoA measurements made from these scattered and reflected paths alone are fairly uncorrelated with the true bearing of the sensor. In such worst-case events, the AoA estimate is drawn uniformly from the feasible set:

$$p_{\text{NLOS}}(\hat{\theta}/\underline{X}) = \frac{1}{\pi}, \quad \hat{\theta} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]. \tag{4.3}$$

When there are significant contributions from both the LOS path and the NLOS multipath, the model depends on whether the receiver is capable of resolving these contributions spatially. This resolving capability is also dependent on the whether the sensor (source) transmits a narrowband or wideband signal.

Narrowband multipath model: With narrowband sensor transmissions, the collector receiver is capable of resolving the arriving combination of LOS path and NLOS multipath in the spatial domain only. For collectors with relatively small number of antenna elements, the receiver can only measure the AoA of the strongest arriving path (or strongest superposition of paths). Hence, each receiver produces a single AoA estimate that, depending on the relative strengths of the LOS and NLOS components, will provide an estimate close to or very "far" from the true bearing of the sensor. Accordingly, we model a typical narrowband scenario as follows: Let α be the fraction of collectors that experience significantly stronger reflected and scattered multipath than the LOS component. In this event, the AoA appears to correspond to the LOS path being blocked and hence, is drawn from the worst-case NLOS model in (4.3). In the remaining instances, when the LOS path is strong, the AoA is drawn from one of the LOS models, for instance (4.1), and we arrive at the following narrowband model:

$$p_{\text{narrowband}}(\hat{\theta}/\underline{X}) = \alpha \; p_{\text{NLOS}}(\hat{\theta}/\underline{X}) + (1-\alpha) \; p_{\text{Gaus}}(\hat{\theta}/\underline{X}). \tag{4.4}$$

This model also represents the least favorable multipath environment, where the LOS component is either present and significant, or is blocked (completely absent), and serves as our candidate model for developing the NLOS suppression algorithm. Note, that distributions with heavy-tails like the Laplacian AoA model in (4.2) or a Cauchy model can also be used to model the narrowband setting, where some AoA estimates are "far" from the true sensor bearing. We numerically investigate these alternate models in Section 4.4.2.

Wideband multipath model: On the other hand, with wideband sensor transmissions, the receiver has an additional degree of freedom and can resolve paths in time as well. When the scattered or reflected paths are sufficiently temporally or spatially separated from the LOS path, the collector is capable of resolving both the LOS path and possibly multiple NLOS paths. Using an idea similar to the narrowband model, the multiple AoA estimates produced by the collector are modeled as follows: one corresponds to the LOS path and is drawn according the LOS models, and the remaining estimates are drawn uniformly from the feasible set, similar to the worst-case NLOS model, as they correspond to multipath with no LOS components (equivalent to LOS blockage). Suppose that a collector resolves L paths, then the resulting estimates are generated according to

$$p_{\text{wideband}}(\theta^{\hat{(1)}}, \theta^{\hat{(2)}} \dots \theta^{\hat{(L)}} / \underline{X}) = \; p_{\text{Gaus}}(\theta^{\hat{(1)}} / \underline{X}) \; p_{\text{NLOS}}(\theta^{\hat{(2)}} / \underline{X}) \dots \; p_{\text{NLOS}}(\theta^{\hat{(L)}} / \underline{X}). \tag{4.5}$$

Although the LOS path is resolved by the receiver in the wideband scenario, it is not known which of the L AoA estimates corresponds to the LOS path. Note that in this wideband model, actual blockage of the LOS might occur resulting in none of the resolved paths being close to the true sensor bearing.

4.2 Localization in LOS scenarios

We consider, in this section, the problem of sensor localization in situations where the channels between the sensor and all the collectors have LOS components and suffer spreading only due to local scattering. We present an algorithm for sequential aggregation of the available AoA estimates (generated according to the LOS model in Section 4.1) to produce the sensor location. Each collector performs linear minimum mean squared error (LMMSE) updates on the *prior* sensor location estimate (received from a previous collector) with its own AoA estimate, and passes the updated estimate to the next collector. This process is continued until all the available AoA estimates are aggregated into the estimate. Before we describe the sequential algorithm in Section 4.2.2, we develop the LMMSE updates at each collector and some coordinate transformations needed to communicate estimates between different collectors in Section 4.2.1. The ML method and the Cramer-Rao Lower Bound (CRLB), which are both used to benchmark the performance of our algorithm, are presented in Section 4.2.3. Finally, some properties of this algorithm are provided in Section 4.2.4.

4.2.1 LMMSE Updates & Transformations

The LMMSE update, described next, only requires knowledge of the second order statistics of the sensor location estimate. Therefore, the *priors* received by each collector consists of the latest estimate of the error covariance of the sensor location in addition to the estimate of the location. As each collector measures an AoA, it is convenient to work in the polar coordinates centered at that collector. This choice of polar coordinates has the added advantage of making the LMMSE update optimal under Gaussian measurement error models like the LOS model in Section 4.1. The AoA θ is always measured from the x-axis of the global cartesian system (see Figure 4.1).

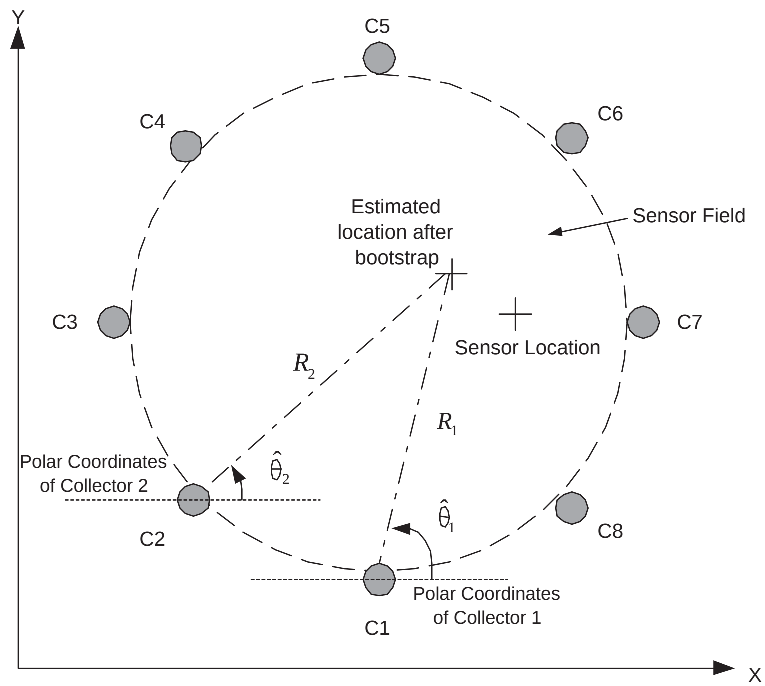


Figure 4.1: Sensor field with 8 collectors on its perimeter: The geometry used to compute the "bootstrap" sensor estimate from AoA estimates of collectors C1 and C2 is shown.

Consider a collector that receives the following prior information: sensor location estimate $\hat{\mu}_{\text{old}} = [\hat{R}_{\text{old}} \quad \hat{\theta}_{\text{old}}]^T$ and error covariance $\hat{\mathbf{\Sigma}}_{\text{old}}$, where

$$\hat{\mathbf{\Sigma}}_{\text{old}} = \begin{pmatrix} \Sigma_{RR}^{(\text{o})} & \Sigma_{R\theta}^{(\text{o})} \\ \Sigma_{R\theta}^{(\text{o})} & \Sigma_{\theta\theta}^{(\text{o})} \end{pmatrix}.$$

The AoA estimate at the collector is $\hat{\theta}_{\text{new}}$ with error variance $\Sigma_{\theta\theta}^{(\text{n})}$. We desire a linear update of the sensor location of the form,

$$\hat{\mu}_{\text{new}} = \hat{\mu}_{\text{old}} + K(\hat{\theta}_{\text{new}} - A\hat{\mu}_{\text{old}}), \tag{4.6}$$

where K is the Kalman gain and $A = [0 \;\; 1]$ (only new AoA estimates are available at the collector). Requiring that the innovation $(\hat{\theta}_{\text{new}} - A\hat{\mu}_{\text{old}})$ be orthogonal to the estimate $\hat{\mu}_{\text{new}}$, we can compute the Kalman gain,

$$K = A\hat{\mathbf{\Sigma}}_{\text{old}} \, (\Sigma_{\theta\theta}^{(\text{n})} + A\hat{\mathbf{\Sigma}}_{\text{old}}A^T)^{-1}. \tag{4.7}$$

Therefore, inserting (4.7) into (4.6), the updated location estimate, $\hat{\mu}_{\text{new}}$, is obtained as

$$\begin{aligned} \hat{R}_{\text{new}} &= \hat{R}_{\text{old}} + \frac{\Sigma_{R\theta}^{(\text{o})}(\hat{\theta}_{\text{new}} - \hat{\theta}_{\text{old}})}{\Sigma_{\theta\theta}^{(\text{o})} + \Sigma_{\theta\theta}^{(\text{n})}}, \\ \hat{\theta}_{\text{new}} &= \frac{\Sigma_{\theta\theta}^{(\text{o})} \; \hat{\theta}_{\text{new}} + \Sigma_{\theta\theta}^{(\text{n})} \; \hat{\theta}_{\text{old}}}{\Sigma_{\theta\theta}^{(\text{o})} + \Sigma_{\theta\theta}^{(\text{n})}}. \end{aligned} \tag{4.8}$$

The update for the covariance matrix can be obtained using (4.6) as

$$\hat{\mathbf{\Sigma}}_{\text{new}}^{-1} = \hat{\mathbf{\Sigma}}_{\text{old}}^{-1} + \begin{pmatrix} 0 & 0 \\ 0 & (\Sigma_{\theta\theta}^{(\text{n})})^{-1} \end{pmatrix} \tag{4.9}$$

and the entries of the updated covariance matrix, therefore, are

$$\begin{aligned} \Sigma_{RR} &= \Sigma_{RR}^{(\text{o})} - \frac{\Sigma_{R\theta}^{(\text{o})} \; \Sigma_{\theta R}^{(\text{o})}}{\Sigma_{\theta\theta}^{(\text{o})} + \Sigma_{\theta\theta}^{(\text{n})}}, \\ \Sigma_{\theta\theta} &= \frac{\Sigma_{\theta\theta}^{(\text{o})} \; \Sigma_{\theta\theta}^{(\text{n})}}{\Sigma_{\theta\theta}^{(\text{o})} + \Sigma_{\theta\theta}^{(\text{n})}}, \\ \Sigma_{R\theta} &= \frac{\Sigma_{\theta\theta}^{(\text{n})} \; \Sigma_{R\theta}^{(\text{o})}}{\Sigma_{\theta\theta}^{(\text{o})} + \Sigma_{\theta\theta}^{(\text{n})}}. \end{aligned} \tag{4.10}$$

When the observations $\hat{\theta}_{\text{new}}$ have Gaussian errors as in the LOS model presented in (4.1), the LMMSE updates in (4.8) are also the optimal minimum mean squared error updates. Under other non-Gaussian error models, the LMMSE updates are the optimal linear updates from the perspective of mean squared error.

In this manner, given priors on the sensor location and new AoA estimates, an updated sensor location estimate $\hat{\mu}_{\text{new}}$ and error covariance $\hat{\mathbf{\Sigma}}_{\text{new}}$ can be produced in a polar coordinate system centered at the collector. This location estimate and error covariance must then be transformed to the global cartesian system (a common frame of reference) to provide the information in a form accessible to the next collector. The next collector then transforms these estimates into its own polar coordinate system. We now describe transformations between the local polar coordinate and global cartesian coordinate systems.

Suppose that the collector is located at $[x_c \; y_c]^T$ in the global cartesian system, and $\hat{\mu} = [R \; \theta]^T$ is the estimated sensor location with error covariance $\hat{\mathbf{\Sigma}}_{\text{pol}}$ in polar coordinates. The estimated sensor location in the cartesian system, $[x_s \; y_s]^T$, is

$$\begin{pmatrix} x_s \\ y_s \end{pmatrix} = \begin{pmatrix} x_c \\ y_c \end{pmatrix} + \begin{pmatrix} R\cos(\theta) \\ R\sin(\theta) \end{pmatrix}. \tag{4.11}$$

Therefore, errors in $[R \; \theta]^T$ can be mapped to errors in $[x_s \; y_s]^T$ as

$$\begin{pmatrix} dx_s \\ dy_s \end{pmatrix} = T_{\text{pol}}(R,\theta) \begin{pmatrix} dR \\ d\theta \end{pmatrix},$$

where

$$T_{\text{pol}}(R,\theta) = \begin{pmatrix} \cos(\theta) & -R\sin(\theta) \\ \sin(\theta) & R\cos(\theta) \end{pmatrix},$$

and the error covariance in the cartesian system is

$$\hat{\mathbf{\Sigma}}_{\text{car}} = T_{\text{pol}}(R,\theta) \; \hat{\mathbf{\Sigma}}_{\text{pol}} \; T_{\text{pol}}(R,\theta)^T. \tag{4.12}$$

Similarly, given the collector location $[x_c \; y_c]^T$ and the sensor estimate $[x_s \; y_s]^T$ in cartesian coordinates, the estimate and covariance can be transformed to the collector's polar coordinates as

$$\begin{aligned} R &= \sqrt{(x_c - x_s)^2 + (y_c - y_s)^2} \\ \theta &= \arctan\left(\frac{y_c - y_s}{x_c - x_s}\right), \end{aligned} \tag{4.13}$$

and

$$\hat{\mathbf{\Sigma}}_{\text{pol}} = T_{\text{car}}(R, \theta) \; \hat{\mathbf{\Sigma}}_{\text{car}} \; T_{\text{car}}(R, \theta)^T, \tag{4.14}$$

where $T_{\text{car}}(R, \theta) = T_{\text{pol}}^{-1}(R, \theta)$.

4.2.2 Sequential Localization Algorithm

We now describe the steps involved in sequentially aggregating collector AoA estimates to produce a location estimate. Let N collectors be located at $\underline{X}_k^c = [x_k \; y_k]^T$ indexed by k, and let the sensor be at $\underline{X}_s = [x_s \; y_s]^T$. Collector k's AoA estimate is $\hat{\theta}_k$ measured from the x-axis of the global cartesian system. Further, the polar coordinate system with collector k at its origin is designated $\mathcal{P}_k$. For ease of exposition, we assume that the collectors are indexed in the order in which their estimates are combined.

The Bootstrap procedure: The algorithm is initialized by considering the first two collectors in the combining order, which are collectors 1 and 2 by convention. Their AoA estimates $\hat{\theta}_1$ and $\hat{\theta}_2$ are used to obtain an initial estimate for the sensor location and error covariance to "bootstrap" the Bayesian algorithm (see Figure

4.1). The estimated range of the sensor from collector 1 is

$$\hat{R}_1 = \frac{(y_2 - y_1)\cos(\hat{\theta}_2) - (x_2 - x_1)\sin(\hat{\theta}_2)}{\sin(\hat{\theta}_1 - \hat{\theta}_2)}. \tag{4.15}$$

Therefore, we can get an initial sensor location estimate $\underline{\hat{X}}_s^{(1)} = [\hat{x}_s \; \hat{y}_s]^T$ in cartesian coordinates from $\underline{\hat{\mu}}^{(1)} = [\hat{R}_1 \; \hat{\theta}_1]^T$ using (4.11). Errors in estimates of sensor range from collector 1 and 2 can be computed from the errors in the AoA estimates as

$$\begin{pmatrix} d\hat{R}_1 \\ d\hat{R}_2 \end{pmatrix} = \frac{1}{\sin(\hat{\theta}_1 - \hat{\theta}_2)} \begin{pmatrix} -\hat{R}_1 \cos(\hat{\theta}_1 - \hat{\theta}_2) & \hat{R}_2 \\ -\hat{R}_1 & \hat{R}_2 \cos(\hat{\theta}_1 - \hat{\theta}_2) \end{pmatrix} \begin{pmatrix} d\hat{\theta}_1 \\ d\hat{\theta}_2 \end{pmatrix}. \tag{4.16}$$

Using (4.16) and independence of the estimates $\hat{\theta}_1$ and $\hat{\theta}_2$, the entries of the initial covariance matrix Σ_{RR}, $\Sigma_{\theta\theta}$, and $\Sigma_{R\theta}$ can be computed in polar coordinates $\mathcal{P}_1$. This initial error covariance $\hat{\mathbf{\Sigma}}_{\text{pol}}^{(1)}$ in $\mathcal{P}_1$ is transformed to the global cartesian coordinates as $\hat{\mathbf{\Sigma}}_{\text{car}}^{(1)}$ using (4.12) and $\underline{\hat{\mu}}^{(1)}$. Note that the initial covariance matrix could have been computed in $\mathcal{P}_2$ instead of $\mathcal{P}_1$, but both would ultimately lead to the same $\hat{\mathbf{\Sigma}}_{\text{car}}^{(1)}$. We term these preceding steps the 'bootstrap procedure'.

Since this bootstrap procedure is used to initialize the algorithm, care must be taken in choosing the collectors to bootstrap in order to avoid 'bad' initial conditions. It is evident from (4.15) and (4.16) that the initial estimates for the range and covariance become unbounded if $(\hat{\theta}_1 - \hat{\theta}_2) \approx n\pi$ due to the $\sin(\hat{\theta}_1 - \hat{\theta}_2)$ term in the denominator of both equations. Geometrically, this condition is equivalent to the direction of arrivals to two different collectors being roughly parallel (or antiparallel), in which case, the point of intersection of the two directions is very far from the collectors and the sensor field. Therefore, the algorithm is always initialized with a pair of collectors with $\hat{\theta}_1 - \hat{\theta}_2$ significantly different from $n\pi$, which also ensures that the initial estimate is mostly within the sensor field. This idea is similar to the concept of dilution of precision in GPS[13] that describes the effect of the satellite configuration on the location accuracy. Thereafter, the collectors can be combined in any random order and effect of this order on performance is studied in Section 4.4.1.

The sequential algorithm for sensor localization has the following steps:

Step 1 (Bootstrap): Estimate the initial location $\underline{\hat{\mu}}^{(1)}$ and error covariance $\hat{\mathbf{\Sigma}}_{\text{pol}}^{(1)}$ in $\mathcal{P}_1$, the polar coordinates of collector 1, using (4.15) and (4.16). Transform the location and error covariance into the global cartesian coordinates as $\underline{\hat{X}}_s^{(1)}$ and $\hat{\mathbf{\Sigma}}_{\text{car}}^{(1)}$ respectively using (4.11) and (4.12). Pass $[\underline{\hat{X}}_s^{(1)}, \hat{\mathbf{\Sigma}}_{\text{car}}^{(1)}]$ as *prior* to the next collector.

Step 2 (Transformation): Let the index of the current collector be k. Transform the priors $[\underline{\hat{X}}_s^{(k-1)}, \hat{\mathbf{\Sigma}}_{\text{car}}^{(k-1)}]$ to $[\underline{\hat{\mu}}^{(k-1)}, \hat{\mathbf{\Sigma}}_{\text{pol}}^{(k-1)}]$ in the local polar coordinate system, $\mathcal{P}_k$, using (4.13) and (4.14).

Step 3 (Aggregation): Update the prior estimates $[\underline{\hat{\mu}}^{(k-1)}, \hat{\mathbf{\Sigma}}_{\text{pol}}^{(k-1)}]$ with the AoA estimate of kth collector, $\hat{\theta}_k$, using the LMMSE procedure in (4.8) and (4.10). Transform updated estimates $[\underline{\hat{\mu}}^{(k)}, \hat{\mathbf{\Sigma}}_{\text{pol}}^{(k)}]$ into estimates in the global cartesian coordinates as $[\underline{\hat{X}}_s^{(k)}, \hat{\mathbf{\Sigma}}_{\text{car}}^{(k)}]$.

Step 4 (Termination): If unaggregated AoA estimates exist, pass priors on to the next unaggregated collector and go to *Step 2*, otherwise output the location estimate and Stop.

Since only a mean and covariance need to be passed on from one collector to the next, this algorithm can be implemented in a completely distributed manner, with each collector needing to know only its own location and orientation. While we consider AoA estimates here, this sequential algorithm is quite general, and can, for example, incorporate probabilistic information on the sensor range obtained from signal strength measurements. Further the complexity of this algorithm grows linearly in the number of collectors. This scalability is required to realize the improvement in localization performance with the number of collectors, details of which are given in Section 4.2.3.

4.2.3 ML Estimate and Cramer-Rao Bound

We now develop the ML estimator and Cramer-Rao bound on the localization performance of the best minimum variance unbiased estimator. For analytical simplicity, we work with the Gaussian AoA model in (4.1), however, corresponding bounds for other models like the Laplacian in (4.2) can be easily derived. For the Gaussian AoA error model, the log-likelihood function (within scale factors and constants) for the observed AoA estimates $\{\hat{\theta}_i\}_{i=1}^{N}$ given the sensor location

$\underline{X} = (x, y)$ is

$$L(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N / \underline{X}) = -\sum_{k=1}^{N} \frac{(\hat{\theta}_k - \theta_k(\underline{X}))^2}{\sigma_k^2}, \tag{4.17}$$

where $\theta_k(X)$ is the true bearing of the sensor from collector k and σ^2 is the AoA estimation error variance at collector k. The ML estimator searches for the location $\underline{X}$ that maximizes the log-likelihood function in (4.17). For small σ_k^2, the cost function in (4.17) can be shown to be approximately concave. We have observed numerically that the cost function has an unique global maxima for the parameter values of interest and a gradient ascent approach produces the ML estimate. The ML estimate is shown to achieve CRLB in Section 4.4.1, thus, validating this gradient ascent approach.

The Fisher information matrix (FIM) $\mathbf{J}(\underline{X})$ for this location estimation is

$$\mathbf{J}(\underline{X}) = \begin{pmatrix} J_{xx} & J_{xy} \\ J_{yx} & J_{yy} \end{pmatrix} = \begin{pmatrix} -E\left(\frac{\partial^2 L(\hat{\underline{\theta}}/\underline{X})}{\partial x^2}\right) & -E\left(\frac{\partial^2 L(\hat{\underline{\theta}}/\underline{X})}{\partial x \partial y}\right) \\ -E\left(\frac{\partial^2 L(\hat{\underline{\theta}}/\underline{X})}{\partial y \partial x}\right) & -E\left(\frac{\partial^2 L(\hat{\underline{\theta}}/\underline{X})}{\partial y^2}\right) \end{pmatrix}, \tag{4.18}$$

where $\hat{\underline{\theta}} = (\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_N)$. Define R_k and θ_k as the range and bearing of sensor at location $\underline{X}$ measured from collector k. Then the entries of the FIM for the Gaussian LOS model in (4.1) are

$$\begin{aligned} J_{xx} &= \sum_{k=1}^{N} \frac{\sin^2(\theta_k)}{\sigma_k^2 R_k^2} \\ J_{yy} &= \sum_{k=1}^{N} \frac{\cos^2(\theta_k)}{\sigma_k^2 R_k^2} \\ J_{xy} &= -\sum_{k=1}^{N} \frac{\cos(\theta_k)\sin(\theta_k)}{\sigma_k^2 R_k^2} \\ J_{yx} &= J_{xy}, \end{aligned} \tag{4.19}$$

and the CRLB is $\mathbf{\Sigma}_{\mathrm{CR}}(\underline{X}) = \mathbf{J}^{-1}(\underline{X})$. The total localization error variance is the sum of the variances along the x and y coordinates. The lower bound on the total localization error is

$$\sigma^2_{\mathrm{tot}}(\underline{X}) \;\geq\; \mathrm{Tr}(\mathbf{\Sigma}_{\mathrm{CR}}(\underline{X})) = \frac{\sum_{k=1}^{N} \frac{1}{R_k^2 \sigma_k^2}}{\sum_{k=1}^{N} \sum_{l=k+1}^{N} \frac{\sin^2(\theta_k - \theta_l)}{R_l^2 R_k^2 \sigma_k^2 \sigma_l^2}}. \tag{4.20}$$

Observe that the CRLB is dependent on the location of the sensor $\underline{X}$. To gain insight into the factors determining the localization error, consider an example of N collectors equally spaced on the perimeter of a disc of radius R (as in Figure 4.1). We simplify the analysis further by computing the bound in (4.20) for a sensor at the center of the disc with the assumption that the AoA error variance is the same for all collectors. In fact, it can be easily verified numerically that the CRLB is maximized at the center of disc, and hence, corresponds to the worst sensor position from a localization standpoint. By realizing that $\theta_k = 2\pi k/N$ for a sensor at the center of the disc, the lower bound on the localization error at the center of the disc is

$$\sigma^2_{\mathrm{tot}} \;\geq\; \frac{NR^2\sigma^2}{\sum_{k=1}^{N} \sum_{l=k+1}^{N} \sin^2(\theta_k - \theta_l)} \approx \frac{2R^2\sigma^2}{\pi^2(N-1)}, \tag{4.21}$$

using the fact that $\sin(\theta) \leq 1$. We note, from (4.21), that the localization error increases linearly with the AoA estimation error variance. Also, the localization error increases with distance from the collector; a given angular error produces increasing location errors at increasing distances from the collector. This, in addition to the observation that the CRLB decreases with distance from the center of the sensor field, points toward the benefits of having at least a few collectors proximal to every part of the sensor field. In other words, AoA estimates from nearby collectors provide the most spatial information. Finally, even with a conservative bound, we observe that the localization error goes down at least as the inverse of the number of collectors providing a method of reducing the localization error in the LOS scenario.

4.2.4 Properties of the Sequential Algorithm

The log-likelihood function in (4.17) is a nonlinear function of the sensor location and the ML estimate is the solution to this non-linear least squares (LS) problem. The solution to this LS problem is asymptotically (in the number of collectors N) consistent and efficient when the observation noise is Gaussian (e.g., LOS AoA error model in (4.1)), and the LS solution is asymptotically consistent even when the noise is non-Gaussian, but with no guarantees on efficiency [71]. The ML estimate can be computed maintaining the asymptotic consistency and efficiency with the following recursive algorithm[71]:

$$\underline{\hat{X}}_n = \underline{\hat{X}}_{n-1} + \mathbf{J}^{-1}(\underline{\hat{X}}_{n-1})\frac{\partial L(\hat{\theta}_n/\underline{\hat{X}}_{n-1})}{\partial \underline{X}}, \quad n = 2, ..., N \tag{4.22}$$

where $\underline{\hat{X}}_n$ is the sensor location after combining the first n collector AoA estimates, $\mathbf{J}(\underline{\hat{X}}_{n-1})$ is the FIM for localization with only the first $n-1$ collectors, and $L(\hat{\theta}_n/\underline{\hat{X}}_{n-1})$ is the contribution of the nth observation to the log-likelihood function in (4.17). As in Section 4.2.3, let R_n and θ_n represent the current location estimate $\hat{X}_{n-1}$ in the polar coordinates centered at collector n. We can now rewrite the recursion for this localization problem as

$$\underline{\hat{X}}_n = \underline{\hat{X}}_{n-1} + \mathbf{J}^{-1}(\underline{\hat{X}}_{n-1}) \begin{pmatrix} R_n \sin(\theta_n) \\ -R_n \cos(\theta_n) \end{pmatrix} (\hat{\theta}_n - \theta_n(\hat{X}_{n-1})), \quad n = 2, ..., N.$$

The above equation has exactly the same form as the LMMSE update in (4.6) with the only difference being that this is an update in the global cartesian coordinates rather than the local polar coordinates of collector n. The hessian-like FIM $\mathbf{J}$ is also updated at each step and new FIM $\mathbf{J}(\underline{X}_n)$ can be computed using (4.19):

$$\begin{aligned} \mathbf{J}(\underline{\hat{X}}_n) &= \mathbf{J}(\underline{\hat{X}}_{n-1}) + \begin{pmatrix} \frac{\sin^2(\theta_n)}{\sigma_n^2 R_n^2} & \frac{-\cos(\theta_n)\sin(\theta_n)}{\sigma_n^2 R_n^2} \\ \frac{-\cos(\theta_n)\sin(\theta_n)}{\sigma_n^2 R_n^2} & \frac{\cos^2(\theta_n)}{\sigma_n^2 R_n^2} \end{pmatrix} \\ &= \mathbf{J}(\underline{\hat{X}}_{n-1}) + (T_{\mathrm{pol}}^{-1}(R_n, \theta_n))^H \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\sigma_n^2} \end{pmatrix} T_{\mathrm{pol}}^{-1}(R_n, \theta_n), \end{aligned} \tag{4.23}$$

where $T_{\mathrm{pol}}(R_n, \theta_n)$ is the unitary matrix in (4.12) used to convert covariance estimates from local polar coordinates to global cartesian coordinates. Recalling that the FIM is the inverse of the covariance matrix, observe that the covariance update in (4.23) is equivalent to the update in the sequential algorithm in (4.9) with appropriate transformations to cartesian coordinates. Thus, the sequential algorithm is an alternate formulation of this recursive ML computation and hence, inherits its properties. However, LS solutions are very sensitive to *outliers* in the observations, and we present a robust algorithm next in Section 4.3.

4.3 Localization in NLOS settings

We now address the narrowband multipath scenario, described in Section 4.1, in which the collector receiver can only resolve the stronger of the contributors between the LOS path and reflected or scattered multipath. In the worst-case, the received multipath at some collectors corresponds to LOS blockage, and the resulting AoA are determined solely by NLOS components that are inconsistent with AoAs seen by collectors with strong LOS paths. The NLOS AoA estimates are "far" from the true sensor bearing and direct application of the sequential algorithm in Section 4.2.2 produces poor results. These NLOS AoA estimates act as *outliers* in the LS algorithm and corrupt the location estimate. Since the collectors experiencing strong NLOS paths are not known a priori, robust techniques are necessary to detect the *outliers* and estimate the sensor location using only the 'good' LOS AoA estimates.

On the other hand, with wideband sensor transmissions, the receiver can resolve both the LOS and NLOS components, and produces multiple possible estimates of the sensor bearing, (see wideband multipath model in Section 4.1). In this event, the algorithm must utilize the LOS AoA estimates, but reject estimates that are strongly influenced by multipath. We present in Section 4.3.2 an algorithm capable of *outlier* suppression to handle both the narrowband and wideband scenarios. However, we focus on the former due to its simplicity and indicate how the algorithm can incorporate multiple AoA estimates at each collector as well. But, before we develop this extension of the sequential algorithm in Section 4.2.2, it is instructive to study the ML method, although the optimal ML approach is computationally intensive. We first develop the ML algorithm in Section 4.3.1 and use the ML algorithm to motivate the *outlier* suppression algorithm.

4.3.1 ML Localization

Let α be the fraction of collectors that experience stronger contributions to the received signal from the NLOS components than the LOS components. We modeled this narrowband scenario in (4.4) as a mixture of the LOS model and the worst-cast NLOS models. Under this model, the log-likelihood function for the sensor location $\underline{X}$ is

$$L(\underline{\hat{\theta}}/\underline{X}) = \sum_{k=1}^{N} \log\left[\frac{1-\alpha}{\sigma\sqrt{2\pi}(1-2Q(\frac{\pi}{2\sigma}))}\exp\left(-\frac{(\hat{\theta}_k - \theta_k(\underline{X}))^2}{2\sigma^2}\right) + \frac{\alpha}{\pi}\right], \quad (4.24)$$

where $\underline{\hat{\theta}}$ is the vector of observed AoAs. Equation (4.24) can be approximated using $\log(e^a + e^b) \approx \min(a, b)$ as

$$L(\underline{\hat{\theta}}/\underline{X}) \approx -\sum_{k=1}^{N} \min\left[(\hat{\theta}_k - \theta_k(\underline{X}))^2, \Theta_{\max}^2\right], \quad (4.25)$$

where

$$\Theta_{\max}^2 = 2\sigma^2 \log\left(\frac{\sqrt{\pi}(1-\alpha)}{\sigma\sqrt{2}\alpha(1-2Q(\frac{\pi}{2\sigma}))}\right). \quad (4.26)$$

Then the ML location estimate $\underline{\hat{X}}$ is

$$\underline{\hat{X}} = \arg\max L(\underline{\hat{\theta}}/\underline{X}) \approx \arg\min \sum_{k=1}^{N} \min\left[(\hat{\theta}_k - \theta_k(\underline{X}))^2, \Theta_{\max}^2\right]. \quad (4.27)$$

The ML algorithm in (4.27) is a standard LS minimization with angular errors bounded by a threshold $\Theta_{\max}$. This cost function ensures that there is no incentive to reducing angular errors larger than the threshold and therefore, outliers or NLOS estimates only have limited effect on the location estimate. As desired, the ML estimator in (4.27) reduces to the ML method for the LOS scenario (see (4.17)) in the absence of outliers ($\alpha \rightarrow 0$), and the threshold $\Theta_{\max} \rightarrow \infty$. To extend the sequential algorithm to the NLOS scenario, we impose this constraint, $\Theta_{\max}$, on the largest observed angular error at each step in the algorithm and describe this modified algorithm in the following section.

4.3.2 Sequential Aggregation with Outlier Suppression

In NLOS scenarios, the localization algorithm must identify the subset of collectors with LOS channels and use these AoA measurements to estimate the sensor location. However, the collectors with NLOS channels (outliers) are unknown and arbitrary in number. The problem of finding the subset of collectors with LOS channels amounts to finding the largest subset of collectors that are in mutual agreement. This search is of exponential complexity as it must explore every subset of the N collectors. Therefore, we resort to a randomization of the algorithm in Section 4.2.2, in which we randomize the choice of the first two collectors used in the bootstrap phase. Thereafter, each subsequent collector's AoA estimate is combined ensuring that angular errors in all the aggregated collectors remain below the threshold $\Theta_{\max}$ from (4.26). This procedure is terminated if no further AoA estimate can be aggregated without causing some angular errors to exceed the threshold. Thus, by bootstrapping with different pairs of collectors, this randomized algorithm produces sensor position estimates corresponding to different subsets of collectors that mutually agree, leading to a list of possible explanations for the observed $\underline{\hat{\theta}}$.

The outlier suppression algorithm is an extension of the sequential algorithm in Section 4.2.2 with 2 key differences: First, at each step, the next collector chosen for aggregation is the one with the smallest angular error. The angular error at a collector is the discrepancy between the bearing of a hypothetical sensor at the current estimated location, $\underline{\hat{X}}_s$, and the AoA estimate measured at that collector, i.e., $e(\hat{\theta}, \underline{\hat{X}}_s) = |\hat{\theta} - \theta(\underline{\hat{X}}_s)|$. The error e can be understood as the empirical estimate of the error in (4.27). The collector with the smallest angular error is the collector whose AoA estimate is most consistent with the current location estimate. Second, after combining a new AoA estimate, the updated sensor location is retained only if the angular errors for the aggregated collectors are below the threshold $\Theta_{\max}$. This ensures that outliers are eliminated using the criterion in (4.27) and simultaneously, the collectors with LOS are used to compute the sensor location using a LS computation. Note that the threshold $\Theta_{\max}$ is dependent on the AoA error variance and could vary from one collector to the next. Here, we assume that the AoA variance and hence the threshold are the same for all collectors to simplify the description. Nevertheless, the variation in AoA variances and the thresholds across the collectors can be easily incorporated into this algorithm.

Sequential aggregation with outlier suppression involves repetitions of the following basic steps with multiple random bootstraps (for brevity, we use the phrase "combine with $\hat{\theta}$" to designate the computation of the new location estimate and error covariance using the appropriate coordinate transformations described in Section 4.2.1):

Step 0 (Initialization): Set the list of collectors already aggregated $\mathcal{A} = \emptyset$ and list of collectors yet to be combined, $\mathcal{C} = \{1, \ldots, N\}$.

Step 1 (Bootstrap): Select a pair of collectors at random, say $\{i, j\}$ that has not been used previously. Compute the initial estimate $\underline{\hat{X}}_s$ and error covariance $\hat{\mathbf{\Sigma}}$ using the bootstrap procedure. Add $\{i, j\}$ to the list of aggregated (inlying) collectors, $\mathcal{A} = \mathcal{A} \cup \{i, j\}$, and remove from the list of remaining collectors $\mathcal{C} = \mathcal{C} - \{i, j\}$.

Step 2 (Angular Error Computation): Compute angular errors $e(\hat{\theta}, \underline{\hat{X}}_s)$ over the remaining collectors $\mathcal{C}$, i.e.,

$$e(\hat{\theta}_i, \underline{\hat{X}}_s) = |\hat{\theta}_i - \theta_i(\underline{\hat{X}}_s)| \ \ \forall\ i \in \mathcal{C}.$$

Step 3 (Candidate Selection and Aggregation): Find collector k with smallest error $e(\hat{\theta}_k, \underline{\hat{X}}_s)$. Combine $\underline{\hat{X}}_s$ with $\hat{\theta}_k$ to obtain new *candidate* estimate $\underline{\hat{X}}'_s$ and covariance $\hat{\mathbf{\Sigma}}'$.

Step 4 (Threshold Verification): If the angular error using the *candidate* location $\underline{\hat{X}}'_s$ is below the threshold for all the aggregated collectors, i.e.,

$$e(\hat{\theta}_l, \underline{\hat{X}}'_s) < \Theta_{\max},\ \forall\ l \in \mathcal{A} \cup \{k\},$$

then retain the *candidate* location and covariance, $\hat{\underline{X}}_s = \hat{\underline{X}}'_s$, and $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}'$. Also add collector k to the list of aggregated collectors, $\mathcal{A} = \mathcal{A} \cup \{k\}$.

Step 5 (Termination): Remove collector k from further consideration $\mathcal{C} = \mathcal{C} - \{k\}$. If there are no remaining collectors ($\mathcal{C} = \emptyset$) then Stop else goto *Step 2*.

The above algorithm is repeated M times with different random initial conditions in order to detect the sensor with a high probability, and each run produces a likely sensor location and a confidence (error covariance) in that estimate. If in a wideband system, a collector resolves two arriving paths, the above algorithm can still be used by introducing a second *virtual* collector at the same location as the original collector and assigning to it the second arriving path. However, care must be taken that both the original and virtual collectors are not part of the same location estimate as it would not be physically possible for two different paths to both contain a strong LOS component. Such a scenario is illustrated with an example in Section 4.4.2.

Choice of M and $\Theta_{\max}$: The performance of the outlier suppression algorithm is determined by the choice of the maximum angular error $\Theta_{\max}$ and the total number of iterations, M. Assuming that the AoA estimation error variance is known at each collector, the threshold $\Theta_{\max}$ can be computed using (4.26), if the fraction of collectors with strong NLOS components α is known. In practice, the largest expected fraction of collectors with strong NLOS components, $\alpha_{\max}$, can be set based on knowledge of area of deployment and the worst case NLOS scenario under which close to optimum performance is desired. The value of $\Theta_{\max}$

thus obtained using (4.26) is conservative for lower levels of multipath scattering, as $\Theta_{\max}$ is monotonically decreasing with α. Although we expect that choosing a smaller $\Theta_{\max}$ might prevent some 'good' LOS AoA estimates from being utilized, it also ensures that NLOS measurements do not corrupt the location estimate.

During the bootstrap phase, two collectors are chosen randomly to seed the sequential algorithm. Success in the sequential estimation depends on selecting two collectors with LOS channels to initiate the algorithm. We have observed from simulations that the algorithm always converges to a solution in the 'vicinity' of the bootstrap location, and therefore an estimate using only LOS estimates is produced if the algorithm is bootstrapped with collectors with LOS channels. Hence, we hypothesize that the bootstrap failure is the predominant cause of failure in the randomization (we verify this numerically in Section 4.4.2). We now try to estimate this probability of failure as a function of M, the number of iterations of the outlier suppression algorithm. Suppose the outlier suppression algorithm is repeated with M *different* seeds, the probability of failure in the bootstrap phase is the probability that at least one collector is an outlier in each of the M attempts. The total number of bootstrap pairs is $P = \binom{N}{2}$ and number of pairs with at least one NLOS collector is $K = P - \binom{\lfloor(1-\alpha)N\rfloor}{2}$, where $\lfloor(1-\alpha)N\rfloor$ is the number of collectors with LOS channels. Then,

$$P(\text{bootstrap failure}) = \begin{cases} \frac{\binom{M}{K}}{\binom{P}{M}} & \text{if } \ M \leq K \\ 0 & \text{if } \ M > K \end{cases} \tag{4.28}$$

When the collectors resolve multiple arriving paths, the above probability of failure computation is modified by replacing the number of collectors by the total number of resolved paths. Depending on the the probability of failure acceptable in the system, (4.28) is used to choose the number of randomizations of the algorithm that are necessary.

A practical issue of interest is to know how the number of randomizations M has to be increased with the total number of collectors N to achieve a fixed probability of bootstrap failure under similar NLOS propagation environments (fixed α). A tight upper bound on the probability of bootstrap failure, presented in the Appendix as (A.2), is given by

$$P(\text{bootstrap failure}) \leq \left(\frac{K}{P}\right)^M .$$

Rearranging the terms, we get an tight upper bound on the possible value of M as

$$M \leq \frac{\log(P(\text{bootstrap failure}))}{\log(\frac{K}{P})},$$

where

$$\frac{K}{P} = 1 - \frac{\binom{\lfloor(1-\alpha)N\rfloor}{2}}{\binom{N}{2}} = 1 - \frac{(\lfloor(1-\alpha)N\rfloor)(\lfloor(1-\alpha)N\rfloor - 1)}{N(N-1)} \approx 1 - (1-\alpha)^2,$$

for large N. Therefore,

$$M \leq \frac{\log(P(\text{bootstrap failure}))}{\log(1-(1-\alpha)^2)}. \tag{4.29}$$

We observe that the probability of bootstrap failure is independent of N. This ensures that the outlier suppression algorithm, which has complexity $O(MN)$, still scales linearly in the number of measurements.

4.4 Numerical Results

We study the performance of the proposed algorithms via Monte-Carlo simulations. The simulation setup is as follows: We consider a circular sensor field of unit radius with N equally spaced collectors along the perimeter. As seen from our analysis in (4.21), the localization error grows linearly with distance from the collector. Therefore, by selecting a field of unit radius, we obtain scale-invariant (only dependent on dimensionless quantities) measures of performance. Let the collectors be located at $[\cos(2\pi(k-1)/N)\ \sin(2\pi(k-1)/N)]^T, k = 1, \ldots, N$. Each collector performs AoA measurement on the signal received at its antenna array and representative AoA estimates are generated according to the models described in Section 4.1. We assume throughout that all the collectors have the same AoA estimation error variance for convenience, although the algorithm does not require this. It is, further, assumed that only a single sensor transmits at any given time. We use the Cramer-Rao bound (4.20) as a performance benchmark, but this bound is dependent on the true location of the sensor. In order to have a fair comparison, we run equal number of iterations on each of 25 candidate sensor locations, and compare the total rms localization error against the average of the CRLB at those 25 locations.

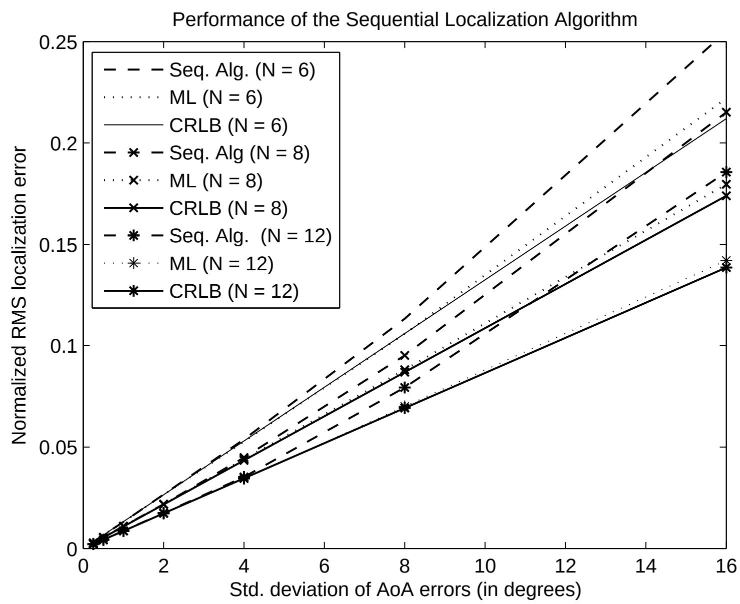


Figure 4.2: Localization performance of the sequential estimation algorithm (in dashed lines) in the LOS scenario for 6,8 and 12 collectors. The performance is compared against the CRLB (in solid lines) and the ML estimator (in dotted lines)

4.4.1 Performance under LOS scenarios

The performance of the sequential algorithm in Section 4.2.2 for different number of the collectors is shown in Figure 4.2. The algorithm achieves the CRLB for small angular estimation errors even for as few as 6 collectors but the performance deteriorates for large AoA error variances (spreads). The optimal ML estimator in Section 4.2.3 can be shown to be approximately convex, and the likelihood function has a unique maxima. The nonlinear coordinate transformations make the sequential algorithm converge to slightly different solutions depending on the order in which the AoA estimates are combined. This effect leaves a performance gap to the ML method, which is significant only at large variances as can be seen from Figure 4.2.

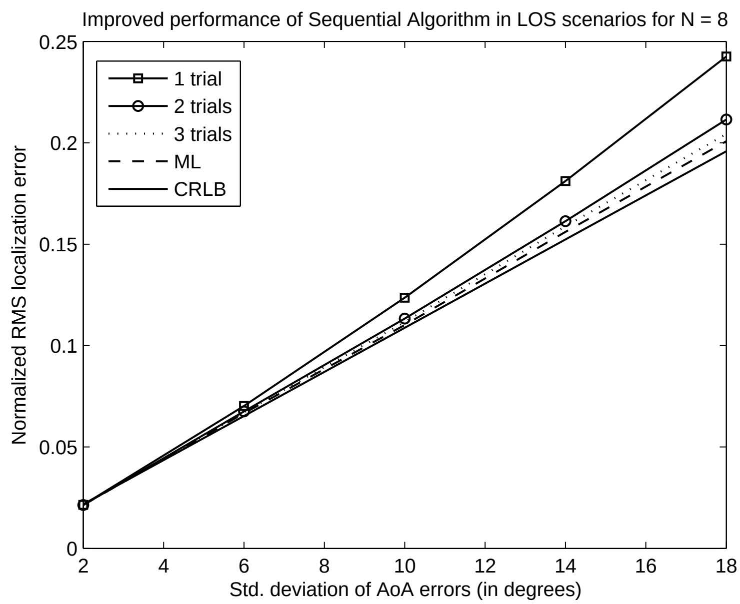


Figure 4.3: Improved performance of the sequential algorithm under LOS settings using multiple random bootstraps (1,2 or 3) for $N = 8$ collectors

The remaining gap to the ML performance can be closed by using multiple random bootstraps and selecting the estimate with smallest ML cost, as shown in Figure 4.3. Therefore, we can also conclude that the combining order for the sequential algorithm has a fairly limited effect on performance. This idea of randomly seeding the sequential algorithm serves as a natural transition to the NLOS localization algorithm in Section 4.3.2. The convergence of the sequential algorithm to the ML solution within three random initializations (bootstraps) in a system with 8 collectors and angular spread of 10° is illustrated with an example in Figure 4.4(Note that a sensor field of radius 500 m is used instead of a unit disc in this illustration). This also supports our earlier finding that the sequential algorithm can provide ML performance in as few as three random initializations.

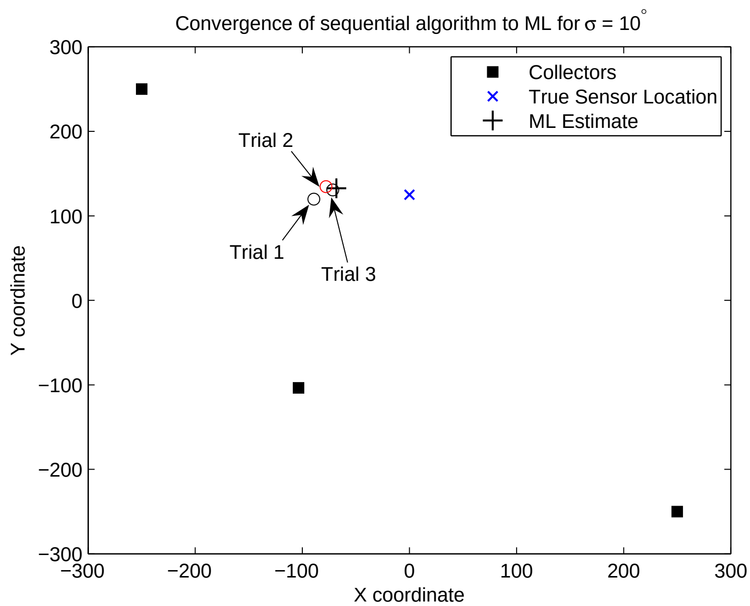


Figure 4.4: Convergence of the sequential algorithm to the ML estimate in a LOS setting with an angular spread of 10° in a system with 8 collectors.

The localization accuracy can also be improved by the addition of collectors to the system. In Figure 4.5, we use a log-log plot to study the dependence of the localization error on the number of collectors N. A linear fit of the simulated data shows that the resolution scales linearly with $N-1$, which agrees with the dependence deduced from the Cramer Rao bound in (4.21). The reason the resolution scales as $N-1$ and not N is due to the fact that at least two collectors are needed to locate a source, and we receive the benefit of error averaging only over the remaining $N-1$ AoA estimates.

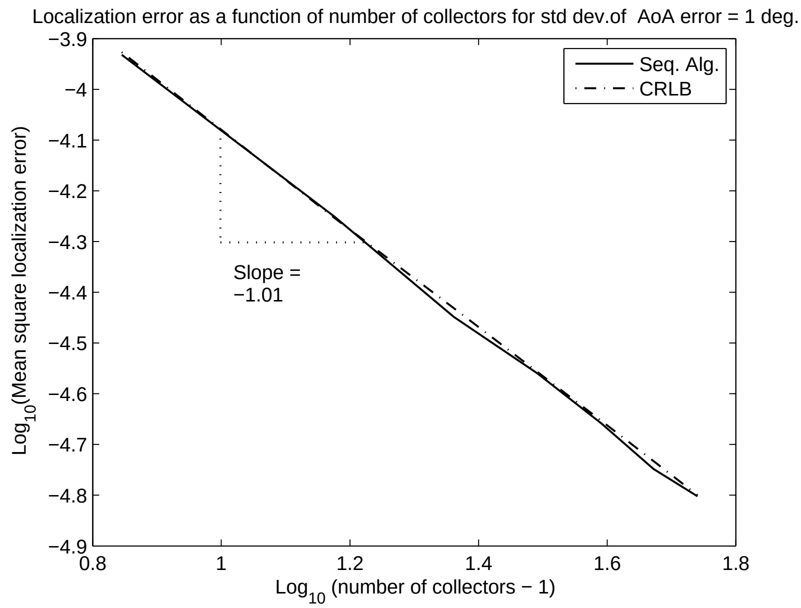


Figure 4.5: Log-Log plot of the mean square localization error against number of collectors N using the sequential algorithm with AoA standard deviation of 1^o. The Cramer-Rao bound is also plotted alongside.

4.4.2 Performance under NLOS scenarios

In this section, we explore the capabilities of the outlier suppression algorithm under the narrowband and wideband multipath models.

Effect of multipath in a narrowband system

We first present numerical results for the narrowband multipath model, where the collectors can only resolve paths spatially and each collector produces a single AoA estimate corresponding to either the LOS path, or the reflected and scattered multipath. In Figure 4.6, the localization error using the proposed outlier suppression algorithm in Section 4.3.2 is compared against the optimal ML estimate in Section 4.3.1 for a system with 8 collectors. For different fractions of outliers, α, the outlier suppression algorithm is run with the threshold $\Theta_{\max}$ chosen according to (4.26), while the number of random bootstraps is chosen to ensure that the probability of bootstrap failure is less than 10^{-3}. This resulted in a choice of the number of random seeds, $M = \{4, 7, 11, 15\}$, for $\alpha = \{0.125, 0.25, 0.375, 0.5\}$ or $\{1, 2, 3, 4\}$ outliers using (4.28). The algorithm puts out multiple solutions, one corresponding to each random initialization. After pruning out the estimates that placed the sensor outside the sensor field boundary, the ML cost function in (4.24) is used to select the most likely estimate. The ML estimate was obtained by brute force minimization of the same cost function. The algorithm performs very close to the optimal ML estimator for the entire range of AoA error variances. However, it is interesting to note in Figure 4.7 that the ML estimate does perform significantly worse compared to the ML estimate using only the good LOS AoA estimates. This additional loss is the cost of identifying the NLOS collectors and as the variance increases, it becomes progressively more difficult to differentiate between the LOS and NLOS AoA estimates.

In practice, since the fraction of outlying collectors, α, is unknown, the algorithm is operated with a threshold chosen using an upper bound on this fraction, which in our simulations is chosen to be $\alpha_{\max} = 0.5$. For this choice of threshold, the localization error is plotted against the ML estimation error for different number ($\{0, 1, 2, 3, 4\}$) of NLOS collectors in Figure 4.8. The outlier suppression algorithm achieves close to ML performance for smaller fractions of outliers even

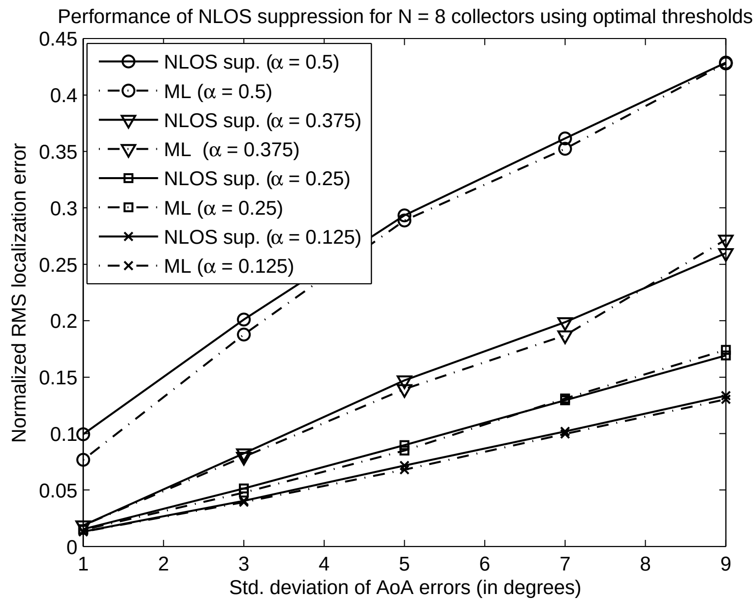


Figure 4.6: Localization performance with NLOS suppression for different fractions, α, of collectors with NLOS channels with optimal thresholds chosen according to (4.26). The location error variance (solid line) is compared against the ML error variance (dash-dotted line).

for this conservative choice of threshold. Thus, we can conclude that this approach is quite insensitive to the exact choice of threshold, $\Theta_{\max}$, which adds to the robustness of this approach. The simulations results also indicate that the algorithm in Section 4.3.2 is approximately ML (AML).

In Figure 4.9, the observed probability of failure of the outlier suppression algorithm is plotted against the expected probability from (4.28) for $N = 8$ collectors. A failure is declared if the location estimate is outside a circle of radius three times the standard deviation of the ML algorithm. When the estimation errors are normally distributed, as is expected from our analysis of the sequential algorithm in Section 4.2.4, the probability of AoA measurement "noise" alone

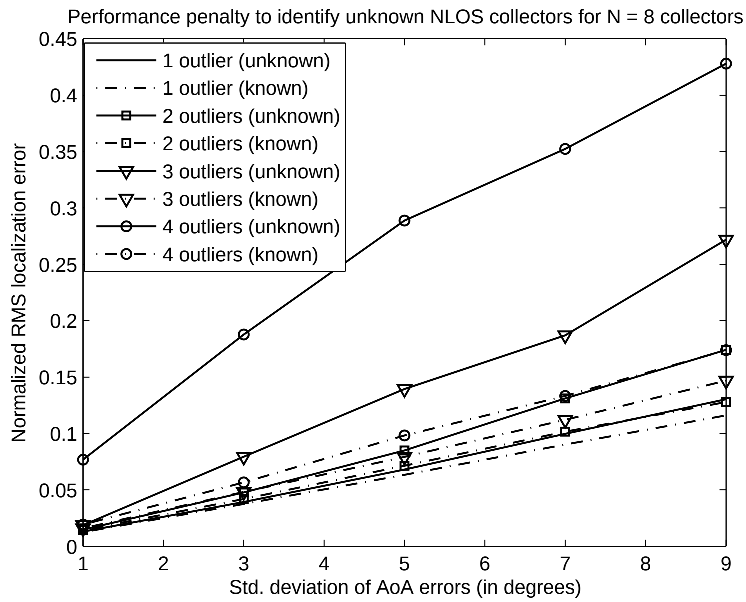


Figure 4.7: The ML localization performance for different number of collectors with NLOS channels for the 2 situations when the collectors with NLOS channels are known or unknown. The gap between the performance when the NLOS collectors are known (dash-dotted line) and are unknown (solid line) represents the performance penalty for outlier identification.

causing the estimate to lie outside this circle is $\exp(-3^2) \approx 10^{-4}$. Hence, failures due to bad bootstraps are significant only when the observed failure probability is of the order of 10^{-3} and above. We observe from the figure that the expected probability of the failure is greater than the observed probability over almost the entire range of interest. However, for all three fractions of outliers, the failure probability plateaus around 3×10^{-3}. This, we believe, is due to the fact that localization errors are not strictly normal in the presence of outliers leading to a slightly higher failure rate due to AoA measurement "noise". We can safely

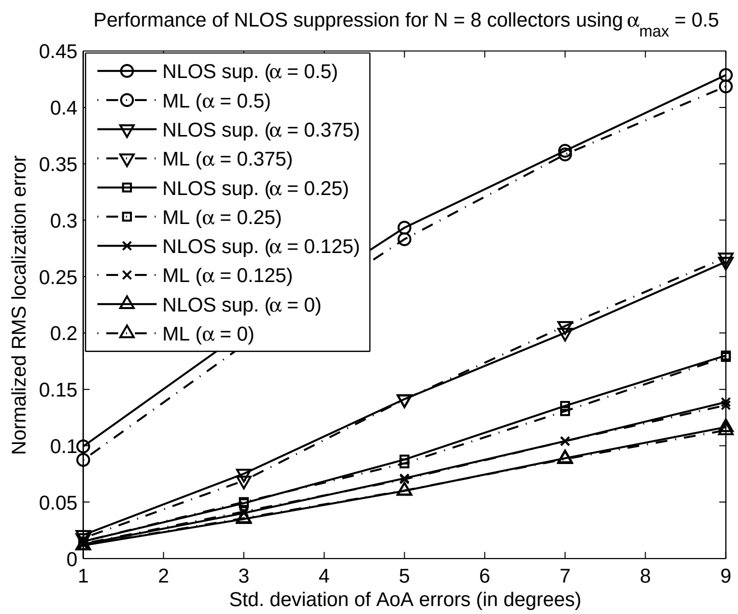


Figure 4.8: Localization performance with NLOS suppression in a realistic scenario with threshold chosen with the maximum expected outlier fraction, $\alpha_{\max} = 0.5$ for different fractions, α, of collectors with NLOS channels. The location error variance (solid line) is compared against the ML error variance (dash-dotted line).

conclude, therefore, that bootstrap failures dominate the failure probability over the entire range of interest.

Effect of alternate narrowband multipath models

Although it is clear that this algorithm is AML for AoA estimates generated using the LOS and narrowband multipath models developed in Section 4.1, it of interest to investigate the sensitivity of the NLOS suppression algorithm to these models. To this end, the algorithm was simulated with two heavy tailed AoA error models, namely the Laplacian (in (4.2)) and Cauchy (the standard deviation

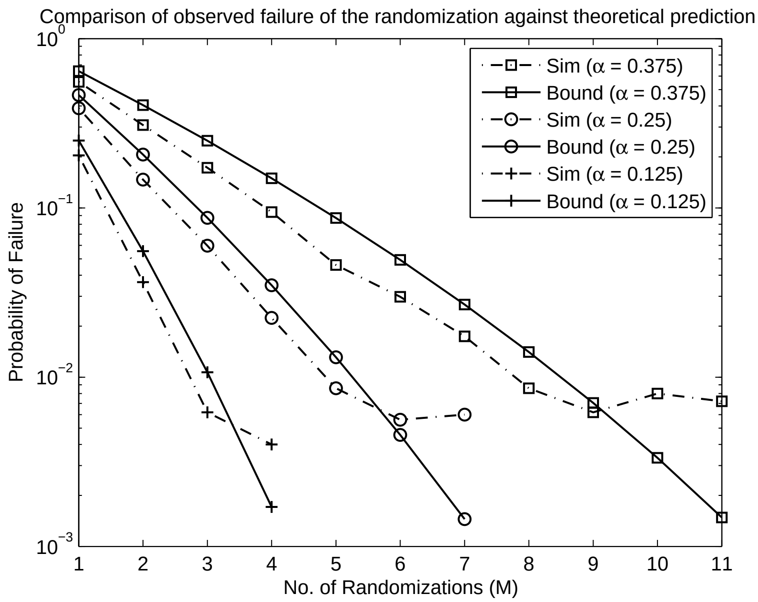


Figure 4.9: The comparison between the observed probability of failure and the expected probability due to bootstrap failures alone for a system with $N = 8$ collectors with different fractions of outliers α. The algorithm fails if the location estimate lies farther than three times the standard deviation of ML estimator.

corresponds to the shape parameter here), for the AoA estimation errors with the same choice of parameters as in the previous example. The resulting performance is shown in Figure 4.10. Under the Laplacian model, which has smaller tails, the algorithm attains the optimal ML performance for that model. But, with a Cauchy model, the heavier tail generates many more outlying AoA estimates leading to larger estimation errors, and the observed performance is equivalent to that of our nominal model in (4.4) for $\alpha = 0.375$.

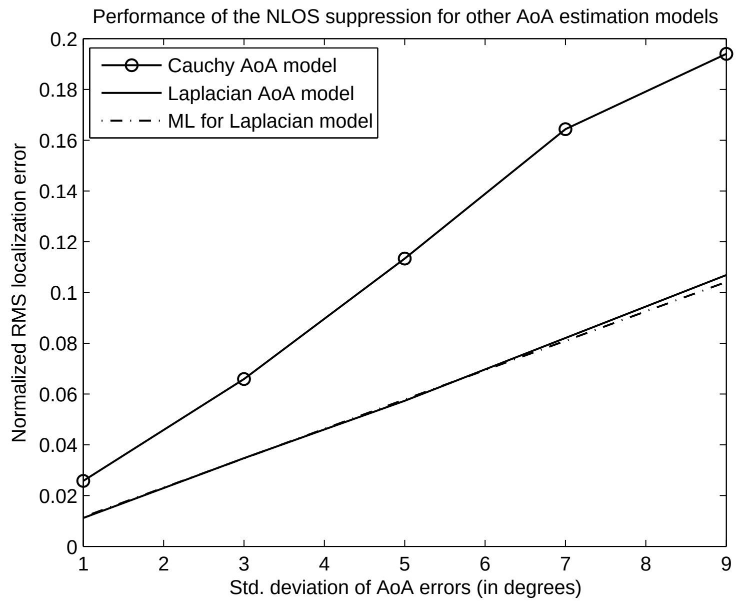


Figure 4.10: Localization performance of NLOS suppression when AoA estimates are obtained from Laplacian and Cauchy models with threshold chosen with the maximum expected outlier fraction, $\alpha_{\max} = 0.5$, as before.

Ray-tracing Illustrations

We now depict the performance of the outlier suppression algorithm in the narrowband and wideband settings with the following two examples. In both examples, we use a virtual point source (shown as a solid circle) model to trace multipath generated by the reflectors; the LOS path from the virtual source to a collector corresponds to the nominal direction of arrival of the reflected signal from the true source. Note that in all the examples, AoA estimation errors of standard deviation $0.5°$ are added to the nominal direction of arrival of the LOS path and the multipath from the virtual source.

Narrowband multipath setting: In figure 4.11, the four collectors (squares) attempt to locate the sensor ('+' sign) in the presence of two reflectors (solid gray lines). One wall blocks the LOS path to the collector C4, causing that collector to only receive multipath reflected from the second wall. The collectors C1 and C2 have LOS AoA estimates, while collector C3 receives both the LOS path and multipath from the virtual source (i.e. reflecting wall). The narrowband receivers generate one AoA estimate each, corresponding to the superposition of all the arriving paths. Thus, collectors C1 and C2 have very reliable LOS AoA estimates, C3 estimates an AoA with a large spatial spread and collectors C4 'sees' an outlying AoA measurement. The estimated sensor locations from multiple runs (crosses) of the outlier suppression algorithm are shown in Figure 4.11. The availability of reliable LOS estimates from C1 and C2 produces good estimates of the sensor location by eliminating the outlying AoA estimate from C4 and also prevents the algorithm from mistaking the virtual sensor to be a real source. Good performance can, therefore, be expected in the narrowband setting even in the presence of LOS blockage if there are sufficiently many collectors with LOS paths to the source. However, there are situations, like when all the collectors experience NLOS propagation, where the narrowband system performs poorly. We elaborate on this issue in the following example.

Wideband multipath setting: As described in Section 4.1, when the sensor transmits a wideband signal, the collector receivers can additionally resolve arriving paths in time, since multipath components suffer delay with respect to

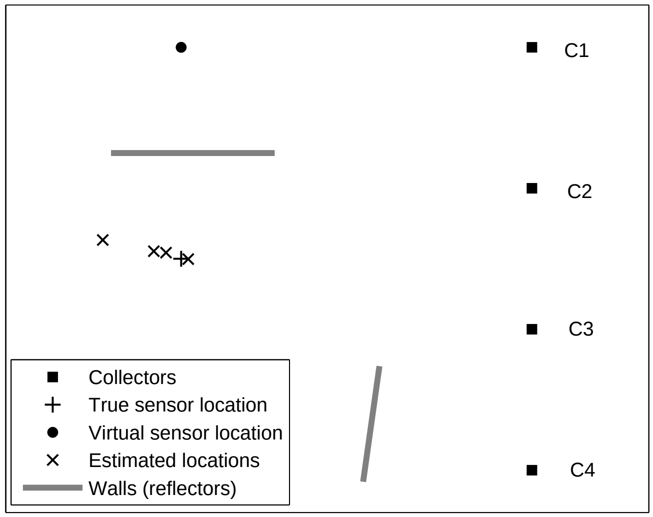
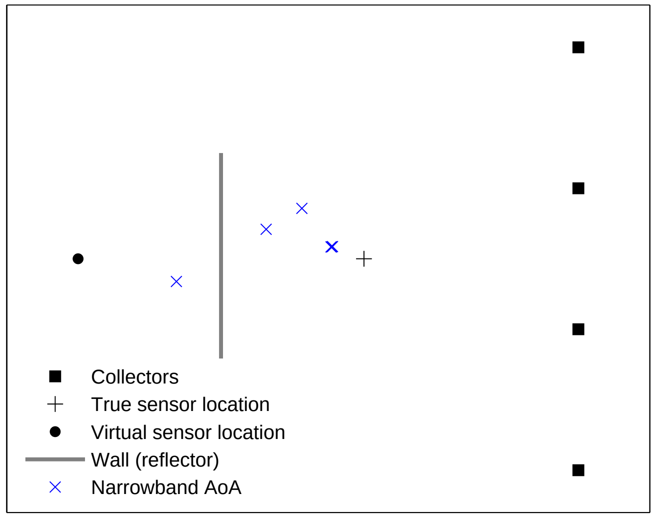


Figure 4.11: The multipath suppression capability of the algorithm in a narrowband setting is illustrated with an example. The collector C4 experiences LOS blockage and receives only reflected multipath from the source. The output of multiple runs of the outlier suppression algorithm with $M = 7$ is shown.

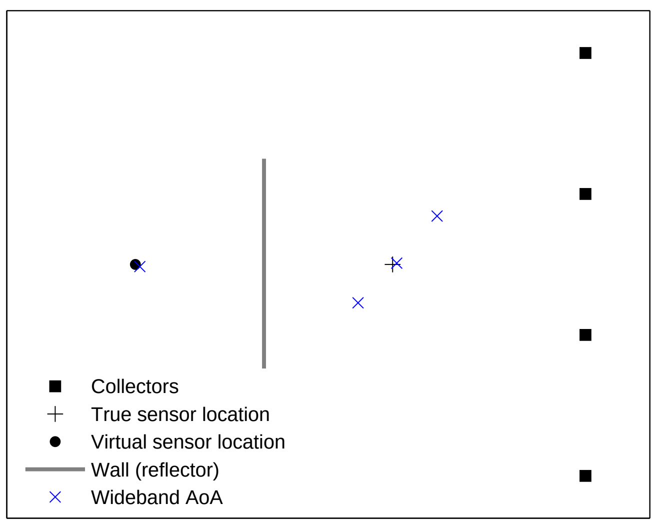
the LOS path due to reflections. This leads to multiple AoA estimates at the collector corresponding to different directions and times of arrival. However, we cannot always reliably conclude that the earliest arriving path is LOS. Instead, we choose to apply our outlier suppression algorithm to all the estimated AoAs and allow the algorithm to eliminate NLOS AoA estimates as outliers. We illustrate this with an example in Figure 4.12. A sensor ('+' sign) is situated between four collectors (squares) and a wall (solid gray line). Under the wideband setting in Figure 4.12(a), each collector generates two AoA estimates, one due to the LOS path and the other due to the reflected NLOS path. On the other hand, with narrowband transmissions in Figure 4.12(b), the receivers estimate the AoA as a power-weighted superposition of the two directions of arrival. The output of multiple runs of the outlier suppression algorithm is plotted for the wideband and narrowband system. In the wideband system, the virtual source location is identified as a likely position in addition to the true sensor location, and the NLOS algorithm is not capable of differentiating between the true sources and virtual sources arising due to correlated multipath. But in practice, the knowledge of the environment can be used to eliminate the infeasible estimates like sensor locations behind the wall. On the other hand, in the narrowband scenario, the source is located in the region between the true and the virtual sources as the multipath in effect increases the spatial spread in the AoA estimates and degrades performance greatly. It is apparent, that the capability to resolve multipath is essential to achieving satisfactory performance in NLOS environments and helps on two counts: First, resolving multiple incoming paths reduces the effective spatial spreading on each path. Second, since multiple estimates are available at each collector, the total number of “good” LOS measurements to estimate the sensor is higher.

4.5 Conclusions

We proposed a sequential algorithm for cooperative localization in *sensor-driven* networks that nearly achieves the CRLB in the presence of LOS between the sensor and the collectors. The localization accuracy (error variance) of this



(a) Narrowband system



(b) Wideband system

Figure 4.12: The working of the localization algorithm in the presence of a virtual source due to a perfect reflector (wall) with wideband and narrowband sensor transmissions. The output of multiple runs of the algorithm is shown.

algorithm under these settings is linear in the spatial spread in the AoA measurements and the size of the coverage area, and inversely proportional to the number of collectors. For instance, in an outdoor environment of area 1 km^2 with local scattering ($3°$ spatial spread, see [81]) at 870 MHz, the obtained localization accuracy using 8 collectors is about 6 m. Correspondingly, with an angular spread of $9°$ in an indoor environment at 5.2 GHz[85], an accuracy of 0.75 m can be obtained in a room of size 30x30 m using just four receivers. In NLOS multipath scenarios, although the outlier suppression algorithm is approximately ML, its performance is heavily dependent on the specific type of environment and the capability of the receivers to resolve the contributions of the LOS path and NLOS multipath in the received signal.

In a narrowband system, where each collector only resolves a superposition of the arriving paths, the algorithm can suppress outlying AoA estimates if there are sufficient number of collectors with reliable LOS estimates. However, the capacity of a wideband system to estimate AoA from the LOS and multipath individually, is vital for good performance in settings where all the collectors experience NLOS propagation. Nevertheless, there is a performance penalty for having to "find" the outliers, which becomes progressively worse as the fraction of outliers in the measurements increases. The proposed algorithms additionally have linear complexity in the number of measurements and are amenable to distributed implementation.

Chapter 5

Conclusions

We now discuss the engineering implications of this work and outline future directions of research.

5.1 Summary

We presented and demonstrated the feasibility of the imaging sensor net architectures for data collection and localization in large-scale sensor networks. We utilized an analogy to imaging, treating the sensors as pixels imaged by a sophisticated collector node, to handle the issue of scale. This resulted in sensors with bare minimum functionality, since the complexity is moved to the collectors. These designs can, thus, provide extremely low-power devices at costs far below the dropping price of today's microsensors, once they are translated into CMOS ICs.

Under this paradigm, there is a shift in emphasis to algorithm design for the sophisticated collectors rather than on algorithms for the sensors themselves. Although this approach is different from the conventional philosophy, we have shown that it can provide benefits such as localization capability in certain applications. These architectures can, indeed, complement prevalent network designs with more sophisticated sensors. A hierarchical network can be envisioned, where the cluster heads gather data from nearby sensors using multihop relay and then communicate this information to the collector using a collector-driven system.

We engineered the RF front-ends at the sensors and leveraged custom RF designs for additional functionality, which, we believe, is a novel approach in the area of sensor networks. Thus, there is potential for application-specific front-end designs for optimized communication links in sensor networks. The fact that the millimeter-wave prototype was built using off-the-shelf components shows that custom front-end designs can be often be built and tested using standard RF blocks. The prototype design also aptly demonstrated the concept of co-design of hardware and algorithms to enable low-power minimalist sensor designs.

However, one important caveat in using RF techniques is that their performance, in practice, is extremely dependent on the propagation characteristics of the environment. Therefore, the system must be appropriately designed and techniques such as those proposed in this dissertation for suppressing effects of multipath (outliers in AoA measurements) need to be developed.

A practical advantage of these architectures is that many system performance parameters can be altered at the collectors independent of the deployed sensor network. For instance, the localization resolution of the collector-driven system can be improved by increasing the bandwidth of SS location code or increasing the transmit power. Similarly, sensors can be added to replace worn out ones or be removed at any time without any bearing on the functioning of the remaining sensors.

By the very nature of imaging sensor nets, the sensor data and location have been fused into one entity, which is very useful in real-time tracking and monitoring applications. This fusion of data and location also presents other possibilities such as location-based retrieval in a collector-driven system, where a selected portion of the sensor field is queried for information.

5.2 Challenges and Future Work

Experiments on the millimeter-wave prototype revealed that *power* is the key resource that determines the feasibility of this system; for instance, we can scale our current system to operate over inter-planetary distances with sufficient transmit power at the collector. Consequently, an important parameter to optimize is the power-efficiency of the system, which can be improved either at the collector or the sensor.

At the collector, more favorable power budgets can be achieved by capturing a larger fraction of the power reflected by the sensors, which can be accomplished by using multiple collectors to record the sensor responses. However, synchronizing all the collectors in time to measure round-trip times or TDOA is both difficult and expensive to achieve in practice. An attractive alternative is for the additional collectors to measure the AoA of the sensor reflection and demodulate the data. Soft estimates of the location and data can be computed using, perhaps, the sequential algorithm proposed in Chapter 4, which can then be incorporated into the estimates of the primary collector. In essence, we have merged the collector-driven and sensor-driven architectures, with the primary collector initiating and providing energy for the sensor transmission.

The $1/R^4$ power fall-off with range of a semi-passive sensors prevents their use in long-range deployments. As mentioned in Section 3.1.1, the additional RF amplification at the sensor (termed an active sensor), can improve the fall-off to $1/R^2$, but at an additional energy cost at the sensors. The collector emits large amounts of power in the beacon, only a small fraction of which gets utilized if there are sensors with data to report. This transmitted energy could be salvaged by sensors that do not have data to send to charge their batteries in a manner similar to energy harvesting [86], and eventually be used for the RF amplification to reach the remote collector. Similarly, the sensor RF circuitry also needs to be turned off, when not in use, to conserve energy, using perhaps a separate wake-up channel.

Alternately, the directivity gain at the sensor can be improved by choosing other antenna designs such as retrodirective arrays [87]. However, the sensor antenna must be (nominally) oriented towards the collector to receive the maximum reflected power, which becomes progressively more challenging with increasing antenna directivity, especially if the sensor is randomly deployed.

The collector-driven system requires LOS between the sensor and the collector to reliably estimate the sensor location, and NLOS multipath causes performance losses. Multipath is particularly significant in the downlink, since at high transmit powers, the sidelobes of the collector transmit antenna can contain a large amount of power as well. Recall that the sensor will reflect the multipath it receives back to the collector as well. Nevertheless, the use of multiple collectors, along with NLOS suppression techniques could help alleviate the effects of multipath on the localization accuracy.

Although we use a SS location code in the prototype, the system is, in principle, capable of operating with other radar waveforms such as a frequency-modulated chirp signal[43], with appropriate modifications. This finding, if confirmed, has a couple of key implications. First, with a chirp signal, the initial range matched filtering in the receiver baseband processing can be performed effectively and rapidly in hardware, and can, thereby, reduce the computation burden on the receiver significantly. Second, a sensor that can reflect a conventional radar waveform can be used as an enhancement in standard imaging platforms. For instance, a SAR remote-sensing system is useful for imaging objects with characteristic response to radar frequencies only. A sensor that has different measurement modalities can either directly modulate the analog measurements or an equivalent bitstream on the radar signal. This side-channel can, thus, potentially provide images of better resolution and higher information content, at the minimal cost of deploying these sensors.

Before the proposed imaging sensor net can be deployed, the sensors need to be integrated with components for sensing, storage and possibly computation. This raises interesting issues of data representation and storage at the sensor, until it is retrieved (or queried for) by the receiver. Moreover, in dense sensor deployments, multiple sensors would detect an event, and the measurements of this event made by nearby sensors would also be highly correlated. Utilization of this spatial correlation between the data transmitted by these nearby sensor for compression (reduction in data sent) or reliability (using the redundancy) is an open problem. In fact, the localization abstraction of the collector-driven system in Chapter 2 could provide a communication infrastructure for very simple sensing modalities such as binary proximity sensors that have been shown to provide reasonable tracking performance in large-scale deployments[88, 89], while utilizing the time correlation between measurements at different sensors as well.

There are also several unresolved theoretical issues in our work on AoA-based localization. Physical layer dependent models need to be extracted in order to understand the effects of the number of antennas and RF front-end components like automatic gain controllers on the localization performance. Such models are also essential to develop techniques for localization of multiple sources that are transmitting simultaneously. The AoA estimation performance is determined by the SNR at the receiver, which varies both with range (due to free-space propagation) and multipath effects, and needs to be accounted for.

Bibliography

[1] F. Zhao and L. Guibas, *Wireless Sensor Networks: An Information Processing Approach.* Morgan Kaufmann, 2004.

[2] A. Cerpa, J. Elson, D. Estrin, L. Girod, M. Hamilton, , and J. Zhao, "Habitat monitoring: Application driver for wireless communications technology." in *ACM SIGCOMM Workshop on Data Communications in Latin America and the Caribbean*, April 2001.

[3] J. Lundquist, D. Cayan, and M. Dettinger, "Meteorology and hydrology in yosemite national park: A sensor network application," in *Proc. Information Processing in Sensor Networks (IPSN)*, April 2003.

[4] S. Meyer and A. Rakotonirainy, "A survey of research on context-aware homes," in *Wearable, Invisible, Context-Aware, Ambient, Pervasive and Ubiquitous Computing (WICAPUC)*, Feb 2003.

[5] S. So, F. Koushanfar, A. Kosterev, and F. Tittel, "Laserspecks: Laser spectroscopic trace gas sensor networks - sensor integration and applications," in *Proc. Information Processing in Sensor Networks (IPSN)*, April 2007.

[6] D. Estrin, L. Girod, G. Pottie, and M. Srivastava, "Instrumenting the world with wireless sensor networks," in *International Conference on Acoustics, Speech, and Signal Processing (ICASSP '01)*, May 2001.

[7] P. Gupta and P. Kumar, "The capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 46, no. 2, pp. 388–404, 2000.

[8] S. Toumpis and A. Goldsmith, "Capacity regions for wireless ad hoc networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 736–748, 2003.

[9] F. Xue, L.-L. Xie, and P. Kumar, "The transport capacity of wireless networks over fading channels," in *Proc. 2004 International Symposium on Information Theory (ISIT 2004)*, 2004, pp. 372–372.

[10] A. Jovicic, P. Viswanath, and S. Kulkarni, "Upper bounds to transport capacity of wireless networks," *IEEE Trans. Inf. Theory*, vol. 50, no. 11, pp. 2555–2565, 2004.

[11] O. Leveque and E. Telatar, "Information theoretic upper bounds on the capacity of large extended ad-hoc wireless networks," in *Proc. 2004 International Symposium on Information Theory (ISIT 2004)*, 2004, pp. 370–370.

[12] R. Min and A. Chandrakasan, "Top five myths about the energy consumption of wireless communication," *ACM Sigmobile Mobile Computing and Communications Review (MC2R)*, Jan 2003.

[13] B. W. Parkinson and J. J. Spilker, *Global Positioning System: Theory and Applications*. American Institute of Aeronautics and Ast, 1996, vol. 1.

[14] J. C. Curlander and R. N. McDonough, *Synthetic Aperture Radar: Systems and Signal Processing*. John Wiley and Sons,Inc., 1991.

[15] "What is radio astronomy," Web, July 2003.

[16] A. R. Koelle, "Short range uhf telemetry system using passive transponders for vehicle id and status information," in *Proc. IEEE Workshop on Automotive Applications of Electronics*, Oct 1988, pp. 34–38.

[17] C. Intanagonwiwat, R. Govindan, and D. Estrin, "Directed diffusion: a scalable and robust communication paradigm for sensor networks," in *Proc. ACM/IEEE International Conference on Mobile Computing and Networking(Mobicom)*, August 2000, pp. 56–57.

[18] J. Liu, D. Petrovic, and F. Zhao, "Multi-step information-directed sensor querying in distributed sensor network," in *Proc. International Conference on Acoustics, Speech and Signal Processing (ICASSP'03)*, vol. 5, Hong Kong, China, April 2003, pp. 145–148.

[19] B. Krishnamachari, Y. Mourtada, and S. Wicker, "The energy robustness tradeoff for routing in wireless sensor networks," in *Proc. IEEE International Conference on Communications(ICC'03)*, vol. 3, Anchorage, AK, May 2003, pp. 1833–7.

[20] W. R. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *HICSS '00: Proc. 33rd Hawaii International Conference on System Sciences*, vol. 8. IEEE Computer Society, 2000, p. 8020.

[21] W. Ye, J. Heidemann, and D. Estrin, "An energy-efficient mac protocol for wireless sensor networks," in *Proc. IEEE Infocom*, vol. 3, June 2002, pp. 1567–1576.

[22] J. Polastre, J. Hill, , and D. Culler, "Versatile low power media access for wireless sensor networks," in *Proc.International conference on Embedded Networked Sensor Systems (Sensys'04)*, New York, NY, 2004, pp. 95–107.

[23] Y. Xu, S. Bien, Y. Mori, J. Heidemann, and D. Estrin, "Topology control protocols to conserve energy in wireless ad hoc networks," CENS, Technical Report 0006, Jan 2003.

[24] N. Li and J. Hou, "Topology control in heterogeneous wireless networks: Problems and solutions," in *Proc. INFOCOM'04*, Mar 2004.

[25] M. Perillo and W. Heinzelman, "Providing application qos through intelligent sensor management," in *Proc. IEEE International Workshop on Sensor Network Protocols and Applications*, May 2003.

[26] S. Kim, R. Fonseca, and D. Culler, "Reliable transfer on wireless sensor networks," in *Proc. IEEE SECON*, 2004.

[27] S. S. Pradhan, J. Kusuma, and K. Ramchandran, "Distributed compression in a dense microsensor network," in *IEEE Signal Process. Mag.*, vol. 19, March 2002, pp. 51–60.

[28] S. S. Pradhan and K. Ramchandran, "Distributed source coding using syndromes (discus):design and construction," in *IEEE Trans. Inf. Theory*, vol. 49, no. 3, March 2003, pp. 51–60.

[29] D. Petrovic, R. Shah, K. Ramchandran, and J. Rabaey, "Data funneling: routing with aggregation and compression for wireless sensor networks," in *IEEE International Workshop on Sensor Network Protocols and Applications*, May 2003, pp. 156 – 162.

[30] A. Boulis, S. Ganeriwal, and M. Srivastava, "Aggregation in sensor networks: and energy-accuracy trade-off," in *IEE International workshop on sensor network protocols and applications*, May 2003, pp. 128–138.

[31] S. D. Servetto, "Distributed signal processing algorithms for the sensor broadcast problem," in *Proc. Annual Conference. on Information Sciences and Systems(CISS)*, March 2003.

[32] M. A. Paskin, C. Guestrin, and J. McFadden, “A robust architecture for distributed inference in sensor networks.” in *Proc. Information Processing in Sensor Networks, (IPSN’05)*, 2005, pp. 55–62.

[33] A. Scaglione and S. D. Servetto., “On the interdependence of routing and data compression in multi-hop sensor networks,” in *Proc. ACM International Conference on Mobile Computing and Networking (Mobicom)*, September 2002, pp. 140–147.

[34] D. Marco, E. J. Duarte-Melo, M. Liu, and D. L. Neuhoff, “On the many-to-one transport capacity of a dense wireless sensor network and the compressibility of its data.” in *Proc. Information Processing in Sensor Networks (IPSN’03)*, 2003.

[35] L. Doherty, K. S. Pister, and L. E. Ghaoui, “Convex position estimation in wireless sensor networks,” in *Proc. IEEE INFOCOM 2001*, vol. 3, Anchorage, Alaska, April 2001, pp. 1655–1663.

[36] A. Savvides, C.-C. Han, and M. B. Srivastava, “Dynamic fine-grained localization in ad-hoc networks of sensors,” in *Proc. Mobile Computing and Networking(MOBICOM 2001)*, July 2001, pp. 166–179.

[37] A. Savvides, H. Park, and M. Srivastava, “The bits and flops of the n-hop multilateration primitive for node localization problems,” in *Proc. First ACM International Workshop on Wireless Sensor Networks and Application*, Atlanta, GA, Sept 2002, pp. 112–121.

[38] D. Niculescu and B. Nath, “Ad hoc positioning system (APS),” in *Proc. IEEE GLOBECOM 2001*, vol. 5, San Antonio,TX, November 2001, pp. 2926–2931.

[39] C. Savarese, J. Rabay, and K. Langendoen, “Robust positioning algorithms for distributed ad-hoc wireless sensor networks,” in *Proc. USENIX Technical Annual Conference*, Monterey, CA, June 2002, pp. 317–327.

[40] N. Bulusu, J. Heidemann, D. Estrin, and T. Tran, “Self-configuring localization systems: Design and experimental evaluation,” *ACM Trans. on Embedded Computing Systems*, vol. 3, no. 1, pp. 24–60, 2004.

[41] N. Priyantha, H. Balakrishnan, E. Demaine, and S. Teller, “Anchor-free distributed localization in sensor networks,” MIT LCS, Tech. Rep. TR-892, Apr. 2003.

[42] S. N. Labs, "Radio frequency tags for tagging, tracking, and locating applications," Webpage, 2005.

[43] D. Hounam and K.-H. Wagel, "A technique for the identification and localization of sar targets using encoding transponders," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 3–7, Jan 2001.

[44] B. G. Colpitts and G. Boiteau, "Harmonic radar transceiver design: Miniature tags for insect tracking," *IEEE Trans. Antennas Propagat.*, p. 28252832, Nov 2004.

[45] R. Stoleru, T. He, J. A. Stankovic, and D. Luebke, "A high-accuracy, low-cost localization system for wireless sensor networks," in *Proc. ACM Sensys 2005*, 2005, pp. 13–26.

[46] "AN/GSQ-187 remote battlefield sensor system (REMBASS)." Military Analysis Network, Federation of American scientists, Feb 2000, http://www.fas.org/ man/dod-101/sys/land/rembass.htm.

[47] J. G. Castao, M. Svensson, and M. Ekström, "Local positioning for wireless sensors based on bluetooth$^{\mathrm{TM}}$," in *Radio and Wireless Conferrence*, IEEE, Ed., September 2004, pp. 195–198. [Online]. Available: http://www.mrtc.mdh.se/index.php?choice=publications&id=0892

[48] D. Blatt and A. Hero III, "Energy-based sensor network source localization via projection onto convex sets," *IEEE Trans. on Signal Processing*, vol. 54, no. 9, pp. 3614–3619, 2006.

[49] M. Rabbat, R. Nowak, and J. Bucklew, "Robust decentralized source localization via averaging," in *Proc. ICASSP'05*, vol. 5, Mar 2005.

[50] J. Liberti and T. Rappaport, *Smart Antennas for Wireless Communication: IS-95 and Third Generation CDMA Applications.* Prentice Hall, 1999.

[51] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. on Signal Processing*, vol. 42, no. 8, pp. 1905–15, Aug 1994.

[52] G. Carter, Ed., *Coherence and Time Delay Estimation.* IEEE Press, 1993.

[53] J. Chen, R. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near-field," *IEEE Trans. Signal Processing*, vol. 50, no. 8, pp. 1843–1854, Aug 2002.

[54] I. K. Thibaut Ajdler, R. Lienhart, and M. Vetterli, “Acoustic source localization in distributed sensor networks,” in *Asilomar Conference on Signals, Systems and Computers*, vol. 3, Nov 2004, pp. 1328–1332.

[55] M. Wax and T. Kailath, “Optimum localization of multiple sources by passive arrays,” *IEEE Trans. on Signal Processing*, vol. 31, no. 5, pp. 1210–1217, Oct 1983.

[56] J. J. Caffery and G. L. Stüber, “Subscriber location in cdma cellular networks,” *IEEE Transactions on Vehicular Technology*, vol. 47, p. 406416, May 1998.

[57] R. A. Maronna, D. R. Martin, and V. J. Yohai, *Robust Statistics: Theory and Methods*. Wiley, 2006.

[58] J. Borras, P. Hatrack, and N. Mandayam, “A decision theoretic framework for nlos identification,” in *Proceedings of IEEE VTC98*, vol. 2, May 1998, pp. 1583–1587.

[59] L. Cong and W. Zhuang, “Non-line-of-sight error mitigation in tdoa mobile location,” in *IEEE Global Telecommunications Conference*, vol. 1, Nov. 2001, pp. 680–684.

[60] S. Venkatraman and J. C. Jr., “Statistical approach to nonline-of-sight bs identification,” in *Proc. 5th International Symposium on Wireless Personal Multimedia Communications*, vol. 1, Honolulu, Hawaii, Oct 2002, pp. 296–300.

[61] S. Srirangarajan and A. H. Tewfik, “Sensor node localization via spatial domain quasi-maximum likelihood estimation,” in *EUSIPCO*, 2006. [Online]. Available: http://www.arehna.di.uoa.gr/Eusipco2006/papers/1568979916.pdf

[62] B. L. Le, K. Ahmed, and H. Tsuji, “Mobile location estimator with nlos mitigation using kalman filtering,” in *IEEE Wireless Communications and Networking*, vol. 3, Mar. 2003, pp. 1969–1973.

[63] P. C. Chen, “A non-line-of-sight error mitigation algorithm in location estimation,” in *IEEE Wireless Communications and Networking Conference*, no. 316-320, Sept. 1999.

[64] R. Casas, A. Marco, J. J. Guerrero, and J. Falc, “Robust estimator for non-line-of-sight error mitigation in indoor localization,” *EURASIP Journal on*

Applied Signal Processing, vol. 2006, pp. Article ID 43 429, 8 pages, 2006, doi:10.1155/ASP/2006/43429.

[65] B. Ananthasubramaniam and U. Madhow, “Virtual radar imaging for sensor networks,” in *Proceedings of the third international symposium on Information processing in sensor networks.* Berkeley, California, USA: ACM Press, New York, April 2004, pp. 294–300.

[66] ——, “On localization performance in imaging sensor nets,” *IEEE. Trans. Signal Processing*, vol. 55, no. 10, pp. 5044–5057, Oct 2007.

[67] D. C. Munson and R. L. Visentin, “A signal processing view of strip-mapping synthetic aperture radar,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 12, pp. 2131–2147, Dec 1989.

[68] J. A. C. Lee, O. Arikan, and D. C.Munson, “Formulation of a general imaging algorithm for high resolution synthetic aperture radar,” in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 4, 1996, pp. 2092–5.

[69] R. E. Blahut, *Theory of Remote Image Formation.* Cambridge University Press, 2004.

[70] D. C. Munson, J. D. O’Brien, and W. K. Jenkins, “A tomographic formulation of spotlight-mode synthetic aperture radar,” *Proc. IEEE*, vol. 71, no. 8, p. 917, 925 1983.

[71] H. V. Poor, *An Introduction to Signal Detection and Estimation*, 2nd ed. Springer, 2005.

[72] F. M. Gardner, “A BPSK/QPSK timing-error detector for sampled receivers,” *IEEE Trans. Comm.*, vol. 34, no. 5, pp. 423–429, May 1986.

[73] ——, “Interpolation in digital modems-part I: Fundamentals,” *IEEE Trans. Comm.*, vol. 41, no. 3, pp. 501–507, March 1993.

[74] L. L. Sharf, *Statistical Signal Processing: Detection, Estimation and Time Series Analysis.* Addison-Wesley, 1991.

[75] J. Proakis, *Digital Communications.* McGraw Hill, 2001.

[76] B. Ananthasubramaniam and U. Madhow, “Collector receiver design for data collection and localization in sensor-driven networks,” in *Proc. Conference on Information Sciences and Systems, (CISS’07)*, March 2007.

[77] L. C. Godara, "Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations," *Proc. of IEEE*, vol. 85, no. 8, pp. 1195–1245, Aug 1997.

[78] R. Raich, J. Goldberg, and H. Messer, "Bearing estimation for a distributed source: Modeling, inherent accuracy limitations, and algorithms," *IEEE Trans. Signal Processing*, vol. 48, pp. 429–441, Feb. 2000.

[79] R. B. Ertel, K. Sowerby, T. S. Rappaport, and J. H. Reed, "Overview of spatial channel models for antenna array communication systems," *IEEE Pers. Commun.*, p. 1022, 1998.

[80] Y. Meng, P. Stoica, and K. M.Wong, "Estimation of the directions of arrival of spatially dispersed signals in array processing," *IEE Proceedings Radar, Sonar and Navigation*, vol. 143, no. 1, pp. 1–9, Feb. 1996.

[81] T. Trump and B. Ottersten, "Estimation of nominal direction of arrival and angular spread using an array of sensors," 1996. [Online]. Available: citeseer.ist.psu.edu/trump96estimation.html

[82] S. Valaee, B. Champagne, and P. Kabal, "Parametric localization of distributed sources," *IEEE Trans. Signal Processing*, vol. 43, no. 9, p. 21442153, Sept. 1995.

[83] Q. H. Spencer, B. D. Jeffs, M. A. Jensen, and A. L. Swindlehurst, "Modeling the statistical time and angle of arrival characteristics of an indoor multipath channel," *Ieee Journal On Selected Areas In Communications*, vol. 18, no. 3, pp. 347–360, Mar. 2000.

[84] B. Rao and K. Hari, "Performance analysis of root-music," *IEEE Trans. on Signal Processing*, vol. 37, no. 12, pp. 1939–1949, Dec 1989.

[85] D. Laurenson, C. Tan, C. Chong, and M. Beach, "Directional measurements and modelling of indoor environments at 5.2 ghz," in *Proc. 2005 European Signal Processing Conference (EUSIPCO'05)*, 2005.

[86] A. Kansal and M. B. Srivastava, "An environmental energy harvesting framework for sensor networks," in *ACM/IEEE Int'l Symposium on Low Power Electronics and Design*, 2003.

[87] R.Y.Miyamoto and T.Itoh, "Retrodirective arrays for wireless communications," *IEEE Microwave Magazine*, vol. 3, no. 1, pp. 71–79, Mar 2002.

[88] N. Shrivastava, R. Mudumbai, U. Madhow, and S. Suri, "Target tracking with binary proximity sensors: Fundamental limits, minimal descriptions, and algorithms," in *Proc. of ACM SenSys*, 2006.

[89] J. Singh, U. Madhow, R. Kumar, S. Suri, and R. Cagley, "Tracking multiple targets using binary proximity sensors," in *Proceedings of the Sixth ACM/IEEE International Conference on Information processing in Sensor Networks (IPSN)*, 2007, pp. 529–538.

Appendix

Appendix A

Bounds on bootstrap failure probability

In this appendix, we derive the bounds on the probability of the bootstrap failure in (4.28). As defined earlier, $P = \binom{N}{2}$ is the total number of bootstrap pairs and $K = P - \binom{\lfloor(1-\alpha)N\rfloor}{2}$ is the number of pairs with at least one NLOS collector, where $\lfloor(1-\alpha)N\rfloor$ is the number of collectors with LOS channels and α is the fraction of collectors with NLOS channels.

$$\begin{aligned} P(\text{bootstrap failure}) &= \frac{\binom{M}{K}}{\binom{P}{M}} \text{ if } M \leq K \\ &= \frac{K(K-1)\ldots(K-M+1)}{P(P-1)\ldots(P-M+1)} \\ &= \left(1+\frac{K-P}{P}\right)\left(1+\frac{K-P}{P-1}\right)\ldots\left(1+\frac{K-P}{P-M+1}\right) \end{aligned} \tag{A.1}$$

In order to obtain an upper bound on the bootstrap failure probability, we replace each ratio of the form $\frac{K-P}{P-k}$ by a larger fraction $\frac{K-P}{P}$ (note that $K < P$ by definition):

$$P(\text{bootstrap failure}) \leq \left(\frac{K}{P}\right)^{M}. \tag{A.2}$$

Similarly, replacing the ratios $\frac{K-P}{P-k}$ by a smaller fraction $\frac{K-P}{P-M+1}$ in (A.1), we get a lower bound,

$$P(\text{bootstrap failure}) \geq \left(\frac{K-M+1}{P-M+1}\right)^{M}. \tag{A.3}$$

The lower and upper bounds clearly converge if $M \ll K < P$, which occurs when N is large. Moreover, these bounds hold only for $M \leq K$, as the probability of bootstrap failure is zero if $M > K$.