SPARSITY BASED LOCALIZATION IN MEDICAL APPLICATIONS:
MEDICAL IMPLANT IN-BODY LOCALIZATION USING WIRELESS
BODY SENSOR NETWORKS, AND INDOOR TRACKING OF
PATIENTS FOR ASSISTIVE HEALTHCARE

BY

MOHAMMAD POURHOMAYOUN

B.S., Isfahan University of Technology, 2002
M.S., Isfahan University of Technology, 2005

DISQUERTATION

Submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate School of
Binghamton University
State University of New York
2013
Accepted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Electrical Engineering
in the Graduate School of
Binghamton University
State University of New York
2013

July 15, 2013

Mark L. Fowler, Chair and Faculty Adviser
Department of Electrical and Computer Engineering, Binghamton University

Zhanpeng Jin, Member
Department of Electrical and Computer Engineering, Binghamton University

Scott Craver, Member
Department of Electrical and Computer Engineering, Binghamton University

Harold W. Lewis III, Outside Examiner
Department of System Science and Industry Engineering, Binghamton University
Abstract

Wireless assistive healthcare technologies, wireless communication medical devices, and body sensor networks for medical diagnostics and therapeutics have attracted increasing attention recently. Modern Healthcare systems use wearable and implantable medical devices to capture and transmit the medical information and vital signs of the patients and control the organs activities.

In this dissertation, we explore the applications of wireless sensor network in biomedical systems including in-body implant localization and indoor tracking of patients for assistive healthcare. The general problem of localization in wireless sensor networks is also considered by modifying the existing methods and designing innovative techniques to reduce the computational complexity and data transmission in sensor networks, as well as designing and implementing new localization methods based on spatial sparsity.

Some unique challenges exist for in-body localization of a medical implant due to the complex nature within the human body, such as the dependency of the propagation velocity on type of the tissues, the dependency of the path loss model on tissue thickness, and the multipath problem caused by signal reflections at organ boundaries. Furthermore,
the safety restrictions on signal power and bandwidth with regard to the protection of human health also make it more difficult to achieve accurate implant location estimation.

In this dissertation, we develop novel and effective methods for location estimation in medical applications including medical implant in-body localization using wireless body sensor networks, and also indoor localization, tracking and fall detection of patients for assistive healthcare. The proposed method estimates the location directly without going through the intermediate stage of Received Signal Strength (RSS) or Time of Arrival (TOA) estimations. In this method, we exploit the spatial sparsity in 3D space and convex optimization theory to achieve very precise results compared to other existing methods. The simulation results demonstrate that the proposed methods are very accurate in location estimation even using a small number of sensors, and a small number of signal samples. Furthermore, the system shows robust operations and high performance in noisy environments (low SNRs). It means that we are able to achieve high localization accuracy even with very low wireless transmitted power which helps to reduce the size of the implantable or wearable device, increase the device’s battery life, and also reduce the risk of interfering with other users of the same band.
To My Dear Laleh

To My Parents

To Dr. Mark L. Fowler
Acknowledgements

I would like to express my deepest gratitude and sincere appreciation to my advisor, Professor Mark L. Fowler for his invaluable guidance, encouragement and continuous support throughout my Ph.D. studies. He is a great individual, advisor and teacher, and has always been a source of inspiration for me.

I would like to thank Professor Zhanpeng Jin for his valuable helps and great guidance and advice concerning my research. I would also like to extend my appreciation to my committee members, Professor Scott Craver and Professor Harold Lewis III for reviewing and evaluating my dissertation.

From the bottom of my heart, I would like to thank my wife and my parents for their never ending love and support. They are the best family one could wish for. Without any of you, this dissertation would have never been possible.
Content

List of Tables ................................................................................................................ x
List of Figures ................................................................................................................ xi
Chapter 1: Introduction ................................................................................................... 1
Chapter 2: Novel Distributed Computation and Data Compression Methods in Wireless Sensor Network ................................................................................................................. 6
  2.1. Emitter Localization in Wireless Sensor Network.................................................. 6
  2.2. Gershgorin Theorem - DPD (GT-DPD) ................................................................. 12
  2.3. Cross Ambiguity Function Sharing - DPD (CS-DPD) .......................................... 17
  2.4. HYBRID CLASSICAL AND DPD (HC-DPD) ..................................................... 30
  2.5. Comparison between DPD and the Proposed Methods ....................................... 35
Chapter 3: A Novel Method: Spatial Sparsity Based Localization in Wireless Sensor Network .......................................................................................................................... 37
  3.1. Localization Based on Spatial Sparsity, and Comparison with Classic and DPD methods ......................................................................................................................... 37
  3.2. Sparsity Approach ............................................................................................... 39
  3.3. Problem Formulation .......................................................................................... 40
  3.4. Comparing Spatial Sparsity Based Method with Classic and DPD methods .......... 44
Chapter 4: Indoor Localization, Tracking and Fall Detection for Assistive Healthcare Based on Spatial Sparsity ........................................................................................................... 54
  4.1. Tracking of Patients and Fall Detection for Assistive Healthcare ....................... 54
  4.2. Problem Formulation .......................................................................................... 56
4.3. Simulation Results for Indoor Localization and Fall Detection .......................... 60

Chapter 5: In-Body Localization Using Wireless Body Sensor Network:
Applications and Challenges.................................................................................. 68
  5.1. Importance and Applications ........................................................................... 68
  5.2. Challenges for In-Body Localization............................................................. 74

Chapter 6: Spatial Sparsity Based Medical Implant In-Body Localization Using
Wireless Body Sensor Network.............................................................................. 77
  6.1. Spatial Sparsity Based Approach for In-Body Localization............................ 77
  6.2. Signal model and parameters ......................................................................... 80
  6.3. Problem Formulation ...................................................................................... 87
  6.4. Reducing Computational Complexity............................................................ 90
  6.5. Simulation Results .......................................................................................... 92

Chapter 7: Conclusion and Future Work................................................................. 102

References.............................................................................................................. 105
List of Tables

Table 2.1: Comparison between DPD and the Proposed Methods ................................ 36

Table 5.1: Path loss parameters: implant to body surface model .............................. 81
List of Figures

Figure 2.1: Classic Two-Stage Localization ............................................................................. 8
Figure 2.2: Placement of the sensors and the emitter position used for simulation............ 17
Figure 2.3: RMS errors for X and Y versus SNR ................................................................. 17
Figure 2.4: Signal sharing for CS-DPD with four sensors .................................................... 20
Figure 2.5: Signal sharing for CS-DPD with five sensors ..................................................... 20
Figure 2.6: Placement of the sensors and the emitter position used for simulation of CS-
DPD with EZW-based compression ..................................................................................... 23
Figure 2.7: RMS errors for X and Y versus bits/element for CS-DPD with EZW-based
compression ......................................................................................................................... 25
Figure 2.8: Singular Values of CAF for two cases: (a) Noise-free signal, (b) Noisy signal 27
Figure 2.9: Placement of the sensors and the emitter position used for simulation of CS-
DPD with SVD-based compression ..................................................................................... 29
Figure 2.10: Simulation results showing RMS error for X and Y when we compress CAF
using SVD-based compression ......................................................................................... 29
Figure 2.11: HC-DPD for four receivers .................................................................................. 33
Figure 2.12: Placement of the sensors and the emitter position used for simulation........... 34
Figure 2.13: RMS errors for X and Y estimates versus SNR ................................................ 34
Figure 3.1: $\ell_2$-norm and $\ell_1$-norm minimization visualization ........................................ 40
Figure 3.2: Placement of the sensors and the emitter position used for simulation............ 45
Figure 3.3: RMS Error for X and Y versus SNR (in dB) for single-emitter single-path
case ..................................................................................................................................... 46
Figure 3.4: Placement of the sensors and the emitter position used for simulation........... 47
Figure 3.5: RMS Error for X and Y (meter) versus SNR (dB) for multipath scenario. ... 48
Figure 3.6: Placement of the sensors and the emitter position used for simulation........... 49
Figure 3.7: RMS Error for X and Y (meter) versus SNR (dB) for estimating the locations of two emitters in (x-y) plane (multi-emitter scenario) .................................................. 50
Figure 3.8: Placement of the sensors and the emitter position used for simulation........... 51
Figure 3.9: RMS Error for X and Y (meter) versus SNR (dB) for estimating the locations of two emitters in (x-y) plane (multi-emitter scenario) in the presence of multipath. ..... 52
Figure 4.1: (a) A typical apartment layout. (b) Four sensors mounted in the corners. (c) A simple case for multipath scenario. The solid line is direct path and dashed lines are reflected paths. ........................................................................................................... 62
Figure 4.2: RMS Error for X and Y (meter) versus Number of sensors.......................... 63
Figure 4.3: (a) True position of the patient (in blue) and the estimated position (in red), (b) Error in positioning for each location in part (a)......................................................... 65
Figure 4.4: True position of the patient (in blue), the estimated position (in red), and the detected position of the fall................................................................. 67
Figure 5.1: Endoscopic Capsules .................................................................................... 70
Figure 5.2: The Calypso System, electromagnetic detector array and console, and implants...................................................................................................................... 73
Figure 6.1: (a) Implant and sensors configuration for three sensors. (b) Sensor array mounted on body surface or mounted inside a wearable jacket. ........................................ 78
Figure 6.2: Minimization over all grid points \( p \) on the surface to find the minimum of \( (L_1+L_2) \) that happens at intersection point \( I \). ................................................................................................. 84
Figure 6.3: An example of tissue configuration from two different views. In both figures, the top surface is the body skin, the middle surface is the boundary between muscle and fat and the third surface is the boundary between muscle and the organ tissue. The red points on the body skin are the sensors and the blue line is the line-of-sight from one of the sensors to the emitter. \( I_1 \) and \( I_2 \) show the intersection points........................................................................ 86
Figure 6.4: Splitting the 3D grid space into 2D grid planes ............................................. 91
Figure 6.5: Sensors location and configuration on the body skin for 4, 8, 12 and 16 sensors. The scale units are in cm................................................................. 94
Figure 6.6: RMS Error for X and Y (in cm) versus Number of sensors for three cases SNR = 0 dB, SNR = 10 dB and SNR = 20 dB................................................... 95
Figure 6.7: (a) True position of the endoscopy capsule (in blue) and the estimated position. (b) Error in positioning for each location in part (a)............................. 97
Figure 6.8: A sample tissue configuration ............................................................... 99

Figure 6.9: (a) A simple pattern for tumor (or implant) movement in X direction, (b) RMS Error for implant location estimation for the movement pattern in (a) in two cases: when the tissue boundary surfaces are perfectly known (red curve) and when there is some uncertainty about the boundary surfaces (blue curve). ........................................ 101
Chapter 1
Introduction

A Wireless Sensor Network (WSN) consists of a collection of receiving sensors which are usually deployed in known locations for the purpose of monitoring and collecting data such as radio signals, sound, temperature, etc. Among numerous applications for WSN such as commercial, industrial, medical, military and environmental applications, passive radio signal localization has attracted enormous attention in the past two decades [1] - [25].

Wireless communication medical devices and wireless body sensor networks for medical diagnostics and therapeutics have also gained increasing attention in the context of quality healthcare. Implantable and wearable devices, such as body sensors, smart pills, pacemakers, insulin pumps, etc. have been playing significant roles in healthcare systems by capturing, transmitting and controlling the vital information of patients. Recent advances in semiconductor technology, analog and digital circuit design, and wireless communication have also enabled designing less expensive and smaller wireless
medical devices which are more convenient to be worn outside or to be implanted inside a human body for special medical interventions.

In the past few years, in-body localization has been widely investigated due to its importance in some applications such as wireless capsule endoscopy and radiation therapy, where physicians need to know and track the accurate position of the wireless medical implant inside the human body [24], [25], [49]-[70].

Wireless Capsule Endoscopy (WCE) has become the preferred technique to diagnose and visualize the human gastrointestinal (GI) tract [49]. Wireless capsule endoscopy has many advantages for physicians and patients compared to other in-body diagnostic methods because it is a non-invasive, highly accurate, and portable procedure [49], [50]. Unlike traditional endoscopic methods, such as gastroscopy and colonoscopy which can only reach the first or last several feet of the gastrointestinal tract, WCE is capable of visualizing the entire GI tract [50], [51]. As a critical component of capsule endoscopic examination, physicians need to know and recognize the precise position of the endoscopic capsule when each image is taken during the process of endoscopy. Thus, accurate capsule localization becomes essential in capsule endoscopy [49], [50] and has been extensively investigated in the past few years [49]-[57].

Tumor tracking in radiation therapy is another important application for in-body localization. Radiation therapy is an effective method to combat cancerous tumors by delivering high doses of ionizing radiation to a tumor to kill or control the growth of malignant cells [58]. Intensity-Modulated Radiation Therapy (IMRT) has greatly increased the ability to deliver an accurate radiation dose to target volumes. Knowing the accurate position of the tumor is a very important issue in radiation therapy because any
small motion in the position of the tumor causes the radiation to be delivered to the surrounding healthy tissues rather than the tumor area, which not only degrades the performance of the treatment because of lack of sufficient dose for the tumor, but also it may cause severe side effects such as secondary cancer [59],[61]. Since the position of the tumor changes during radiation therapy (because of patient movements, respiration, gastro-intestinal, cardiac system and circulation), *real-time* tumor tracking is necessary in radiation therapy treatments in order to deliver an accurate amount of radiation to the tumor without damaging the surrounding healthy tissues [59], [61]. Various techniques have been proposed in the literature for tumor tracking during the radiation therapy process based on implanting several wired or wireless devices (called beacons) inside or in the vicinity of the tumor [59]-[70].

In this dissertation, we explore the applications of wireless sensor network in biomedical systems including in-body implant localization, and also indoor tracking of patients for assistive healthcare. The general problem of localization in wireless sensor network is also considered by modifying the available methods including designing innovative techniques to reduce the computational complexity and data transmission, as well as designing and implementing new methods based on spatial sparsity.

Some unique challenges exist for in-body localization of a medical implant due to the complex nature within the human body, such as the dependency of the propagation velocity on type of the tissues, the dependency of the parameters of the path loss model on tissue thickness (deep or near surface), and the multipath problem caused by signal reflections at organ boundaries. Furthermore, the safety restrictions on signal power and
bandwidth with regard to the protection of human health also make it more difficult to achieve accurate implant location estimation.

The traditional localization methods use two stages to estimate the position. In the first stage, a set of location-dependent parameters (such as time-difference-of-arrival (TDOA), frequency-difference-of-arrival (FDOA), or received-signal-strength (RSS)) are estimated from the received signals. Then in a second stage, these estimated parameters are used in statistical processing to locate the signal source [2]. However, the two-stage method is not necessarily optimal because in the first stage of these methods, the location-dependent parameters are obtained by ignoring the fact that all measurements should be consistent with a single emitter location [12], [13], [14].

In this research, we develop novel and effective methods for location estimation in medical applications including medical implant in-body localization using wireless body sensor networks, and also indoor localization, tracking and fall detection of patients for assistive healthcare. The proposed method estimates the location directly without going through the intermediate stage of Received Signal Strength (RSS) or Time of Arrival (TOA) estimations. In this method, we exploit the spatial sparsity in 3D space and convex optimization theory to achieve very precise results compared to other existing methods. The simulation results demonstrate that the proposed methods are very accurate in location estimation even using a small number of sensors, and a small number of signal samples. Furthermore, the system shows robust operations and high performance in noisy environments (low SNRs). It means that we are able to achieve high localization accuracy even with very low wireless transmitted power which helps to reduce the size of the...
implantable or wearable device, increase the device’s battery life, and also reduce the risk of interfering with other users of the same band.

This dissertation is organized as follows: In chapter 2, we review the main localization methods, and mention the advantages and disadvantages of each method. Then, we propose innovative techniques (including three main methods) to improve the performance of the existing methods using distributed computation algorithms, data compression techniques and etc. In chapter 3, we develop a novel one-stage localization method based on spatial sparsity in general case, and then compare the performance of the proposed method to other common methods in different scenarios. In chapter 4, we develop and apply the proposed sparsity based method for indoor localization, patient tracking and fall detection in assistive healthcare systems, and evaluate the performance using several computer simulations. In chapter 5, we talk about the importance and applications of in-body localization, and then we review the most important challenges existing in medical implant in-body localization. In chapter 6, we develop a novel tissue-adaptive method for in-body localization based on spatial sparsity, using both received signal strength and time of arrival. We also propose some techniques for computational complexity reduction. Then, we evaluate the performance of the proposed method using several sets of simulation for 2D and 3D in-body positioning. Finally, chapter 7 presents the conclusion and some directions for future works.
Chapter 2

Novel Distributed Computation and Data Compression Methods in Wireless Sensor Network

2.1. Emitter Localization in Wireless Sensor Network

Emitter Localization has been a long-standing and important issue in the areas of signal processing and wireless sensor networks that has raised increasing attention in the past decade. There are several efficient methods to estimate the emitter location based on measuring one or more position-dependent parameters of the received signals such as time-of-arrival (TOA), received signal strength (RSS), frequency-difference-of-arrival (FDOA), time-difference-of-arrival (TDOA), angle-of-arrival (AOA) and etc. In the following, we overview the two main approaches for TDOA/FDOA based emitter localization, and mention the advantages and disadvantages of each method. Then, we propose novel techniques to improve the performance of the existing methods using distributed computation algorithms, data compression techniques, etc.
2.1.1. Classic Two-Stage Localization Methods

One of the most accurate and common methods for passive radio signal geolocation is based on exploiting the time-difference-of-arrival (TDOA) and frequency-difference-of-arrival (FDOA) between received signals. The classical approach to this method uses two stages to estimate the signal position. In the first stage, TDOAs and FDOAs are estimated from the cross-correlation of signals received by several pairs of sensors; this is done by computing the cross ambiguity function (CAF), and finding the peak of its magnitude surface [1]. Then these TDOA/FDOA estimates are used in statistical processing to locate the emitter [2]. However, a challenge in such methods is the need to share large amounts of signal data between paired sensors prior to computation of the CAF for each pair, and has recently been addressed in [3], [4].

Suppose that an RF emitter sends out signal \( s(t) \) that is then intercepted by \( L \) moving\(^1\) sensors. The equivalent lowpass (ELP) signal (also called the complex envelope) of the received RF signal at the \( l \)th sensor is given by

\[
\hat{s}_l(t) = \alpha_l \hat{s}(t - \tau_l) e^{j2\pi ft}
\]

(2.1)

where \( \hat{s}(t) \) is the ELP signal of the transmitted signal \( s(t) \), \( f_l \) is the Doppler shift, \( \tau_l \) is the delay, and \( \alpha_l \) is a complex amplitude; this is the commonly used narrowband model [7] for RF signal reception in the presence of relative motion between transmitter and receiver. Using a subscript \( r \) to indicate the noisy received signal we have

\[
\hat{s}_r(t) = \alpha_l \hat{s}(t - \tau_l) e^{j2\pi ft} + w_f(t)
\]

(2.2)

\(^{1}\) In medical applications, the sensors are usually stationary. But, in general localization problems, the sensors can be moving. In this section, we consider the general case in our formulations.
where \( w(t) \) is the ELP signal of the noise. Now, suppose that two sensors \( m \) and \( n \) receive the ELP signals \( \hat{s}_m(t) \) and \( \hat{s}_n(t) \), respectively. The estimate for TDOA/FDOA can be obtained by finding the peak of the magnitude of the CAF; the CAF of the \( m \)th and \( n \)th signals is given by

\[
CAF_{mn}(\tau, \omega) = \int_{-\infty}^{\infty} \hat{s}_m(t)\hat{s}_n^*(t - \tau)e^{-j\omega t} dt.
\] (2.3)

Given TDOA/FDOA estimates made from several pairs of signals, classical TDOA/FDOA location methods statistically process them to arrive at a geolocation in an x-y plane (or more generally, in x-y-z space). Figure 2.1 shows the first and second stages of the classic TDOA/FDOA based emitter localization.

Figure 2.1: Classic Two-Stage Localization: (a) first stage; signal sharing, computing the CAF magnitude, and estimating the TDOA/FDOA, (b) second stage; transmitting the estimated TDOA/FDOA to a common (fusion) point to estimate the location.

We have to notice that the two-stage method is not necessarily optimal because in the first stage of these methods, the TDOA and FDOA estimates are obtained by ignoring the fact that all measurements should be consistent with a single emitter location [12], [13], [14]. In other words, although each TDOA/FDOA estimation can be optimal in the first
stage, and the second stage is also optimal (given the results of the first stage), the whole two-stage method is not optimal.

2.1.2. Direct Position Determination (DPD)

Recently, some new TDOA/FDOA-based methods have been proposed that use only one stage [12], [13], [14], [15], providing improved accuracy. Weiss and Amar [12], [13], [14], introduced a single-stage localization method called Direct Position Determination (DPD). In independent work, Kay and Vankayalapati [15] developed an identical estimator as part of a generalized likelihood ratio (GLR) detector based on all the received signals. The DPD method centrally processes all the received signals to computes the CAF between all pairings of signals in order to form the so-called Cross Ambiguity Matrices (CAM), whose largest eigenvalues are computed and used to estimate the location. Neither [14] nor [15] addressed the issues of computation and data transmission for DPD.

The remainder of this section provides some needed details about the single-stage localization method DPD and its formulation as developed in [14]. Assume that each sensor collects $N$ time samples sampled with sampling frequency $F_s = 1/T_s$. Then, as given in [14] the received sampled signal can be written as

$$\hat{s}_{rl} = \alpha_l W_l D_l \hat{s} + w_l$$

(2.4)

$$W_l \triangleq \text{diag} \{ e^{j2\pi f_{l1}}, e^{j2\pi f_{l2}}, \ldots, e^{j2\pi f_{ln}} \}$$

$$\hat{s}_n \triangleq [\hat{s}_n(t_1), \hat{s}_n(t_2), \ldots, \hat{s}_n(t_N)]^T$$

$$w_l \triangleq [w_l(t_1), w_l(t_2), \ldots, w_l(t_N)]^T$$
where \( \hat{s}_n = [\hat{s}_n(t_1), \hat{s}_n(t_2), \ldots, \hat{s}_n(t_N)]^T \) is a vector of the \( N \) samples of the received signal at \( l \)th sensor, \( \hat{s} \) is a vector of the \( N \) samples of the transmitted signal, \( w_r \) is the noise vector, \( f_l \) is the Doppler shift and \( D_l \) is the time sample shift operator by \( n_l = (\tau_l / T_s) \) samples. We can write \( D_l = D^n \) where \( D \) is an \( N \times N \) permutation matrix defined as \[ (D)_{ij} = \begin{cases} 1 & \text{if } i = j + 1, \quad (D)_{4,0,1,1} = 1 \text{ and } (D)_{4,0} = 0 \text{ otherwise.} \end{cases} \]

According to [14] and [15], the estimated transmitter’s position in TDOA/FDOA-based one-stage method is found as follows. Let \( \{p_i\} \) be grid points of possible emitter locations in the x-y plane (or more generally in an x-y-z space). For each grid point form the matrix

\[
V_{p_i} = [D_1^H W_1^H \hat{s}_r, D_2^H W_2^H \hat{s}_s, \ldots, D_L^H W_L^H \hat{s}_r],
\]

where the delay and Doppler operators correspond to the path between the grid point and the respective sensors 1 to \( L \). For each grid point, then form the \( L \times L \) matrix \( Q_{p_i} = V_{p_i}^H V_{p_i} \) and find its largest eigenvalue, which gives a set of \( G \) largest eigenvalues. The location estimate is the grid point that maximizes the largest eigenvalue, that is

\[
\hat{p} = \arg \max_{p_i, i \in [1, 2, 3, \ldots, G]} \{\lambda_{\max}(Q_{p_i})\}
\]

As shown in [14] and [15], the \( (k,l) \)th element of the matrix \( Q \) is the value of the CAF between the signals received by sensor \( k \) and sensor \( l \) and that is why in [14], this matrix is called the Cross Ambiguity Matrix (CAM). The elements of the CAM are then given by
\[ [Q_{p_i}]_{ij} = [V^H V]_{ij} \]
\[ = \hat{s}_{rk}^H W_k D_k D_l^H W_l^H \hat{s}_{rl} \]
\[ = [\text{CAF}_{kl}]_{(\tau_i, \omega_j)} \]  \hspace{1cm} (2.7)

where \( \tau_i \) and \( \omega_j \) are the corresponding TDOA and FDOA between sensors \( k \) and \( l \) with emitter hypothesized as being located at position \( p_i \). Note that the diagonal elements in the CAM \( (k = l) \) are

\[ [Q_{p_i}]_{kk} = [V^H V]_{kk} \]
\[ = \hat{s}_{rk}^H W_k D_k D_k^H W_k^H \hat{s}_{rk} \]
\[ = \| \hat{s}_{rk} \|^2, \]  \hspace{1cm} (2.8)

which is equal to the energy of the received signals or, equivalently, the peak of the auto ambiguity function (AAF) at the origin.

Thus, DPD computes the CAF between all possible pairs of signals (including self-pairing) to populate the CAM. This requires mapping the CAF to the X-Y plane in the same way as used in [5].

There are some issues when DPD is applied in its originally stated formulation. The main problem is the large amount of computations required in this method, which should be done at a single node (sensor). Thus, a decentralized computational approach that strives to spread the computations more uniformly across the sensors is desirable.

In the following, we propose innovative technologies to optimize DPD algorithm by developing distributed computation methods and data compression techniques. We have developed three different methods aimed at addressing DPD problems (to varying degrees). The results have been also presented in several publications [6], [8], [9], [11], [16], [17].
2.2. Gershgorin Theorem - DPD (GT-DPD)

In this section, we develop a method to distribute the computation and reduce the computational complexity at any one sensor based on eigenvalue approximation. This then increases the flexibility and feasibility of the DPD method.

**Definition** (Gershgorin’s disc) [26]: Assume that $A$ is an $n \times n$ complex-valued matrix with entries $a_{ij}$ and $p_i = \sum_{j \neq i} |a_{ij}|$ is the summation of the absolute values of all non-diagonal elements of the $i^{th}$ row. Then, the set $D_i = \{ z \in \mathbb{C} : |z - a_{ii}| \leq p_i \}$ is called the $i^{th}$ Gershgorin’s disc of $A$. This disc contains the interior and boundary points of a circle with radius of $p_i$ and centered at $a_{ii}$ in complex plane.

**Theorem 1** (Gershgorin’s Theorem) [26]: Every eigenvalue of matrix $A = [a_{ij}] \in \mathbb{C}^{n \times n}$ lies within at least one of the Gershgorin discs. In other words, every eigenvalue $\lambda$ of matrix $A$ satisfies

$$\forall \lambda, \exists i \quad |\lambda - a_{ii}| \leq p_i ; \quad p_i = \sum_{j \neq i} |a_{ij}|. \quad (2.9)$$

**Theorem 2** [26]: Assume that $A$ is an $n \times n$ complex valued matrix with entries $a_{ij}$, $R_i = \sum_i |a_{ij}|$ is the summation of the absolute values of all elements in the $i^{th}$ row, and $T_j = \sum_j |a_{ij}|$ is the summation of the absolute values of all elements in the $j^{th}$ column. Let $R = \max_i R_i$ and $T = \max_j T_j$. Then, the absolute value of each eigenvalue $\lambda$ of matrix $A$ satisfies

$$\forall \lambda, \quad |\lambda| \leq \min(R,T). \quad (2.10)$$
As mentioned before, in TDOA/FDOA-based DPD, the \((i,j)^{th}\) element of the cross ambiguity matrix (CAM) or \(Q\) in equation (2.7) is the value of CAF between the signals received by sensors \(i\) and sensor \(j\), and we name it as \(CAF_{ij}\). The emitter location is estimated by computing the maximum eigenvalues of the CAM (or \(Q\)) at each grid point. Since the CAM is Hermitian and positive definite, the eigenvalues of CAM are real and positive. Moreover, since CAM is a Hermitian matrix, we have, \(R=T\) in Theorem 2 and consequently, \(\min(R, T) = R\). Thus, for the CAM, the inequality in (2.10) can be replaced by

\[
\forall \lambda, \quad \lambda \leq \max_i (CAF_i)
\]

and

\[
\hat{\lambda}_{\text{max}} \triangleq \max_i (CAF_i),
\]

where \(\hat{\lambda}_{\text{max}}\) is the upper bound on eigenvalues of the CAM.

Suppose that we have \(L\) receiving sensors and each one of them broadcasts its received signal to all other sensors in the sensor network. Then, each sensor \(i\) is able to compute all \(CAF_{ij}\) for \(j = 1 \ldots L\) and consequently, it is able to compute \(CAF_i = \sum_{j=1}^{L} |CAF_{ij}|\). Now, if we approximate the largest eigenvalue of CAM by the upper bound on the eigenvalues \((\lambda_{\text{max}} \approx \hat{\lambda}_{\text{max}})\), then the location estimation will be determined by the point having the largest \(\hat{\lambda}_{\max}\).

Here is the outline for the GT-DPD method:
1. Each sensor broadcasts its received signal in the sensor network.

2. Each sensor $i$ computes $CAF_{ij}$'s in TDOA/FDOA plane and then maps them from TDOA-FDOA plane to X-Y (emitter position) plane (as is done in [5]). The mapping will be done very easily knowing the position and velocity of the sensors and also the grid point position.

3. Each sensor $i$ computes $CAF_i = \sum_{j=1}^{L} |CAF_{ij}|$ by adding up the $CAF_{ij}$'s and then finds the peak of $CAF_i$ (named $CAF_{i,peak}$) and its location $(x_{i,peak}, y_{i,peak})$ and then transfers the three numbers $x_{i,peak}, y_{i,peak}$ and $CAF_{i,peak}$ to a common sensor (or to all other sensors since there are just three numbers and there is no communication load to transfer them). Note that it is in this step that the Gershgorin Theorem is exploited.

4. According to the (2.12) and (2.6), the emitter location estimated is taken as the $(x_{i,peak}, y_{i,peak})$ corresponding to the largest $CAF_{i,peak}$ over all $i$.

Note that in the original DPD method, we need to re-compute and form the CAM for each grid point and find the largest eigenvalue of that matrix each time, which requires a large amount of computation, especially when the number of receiving sensors is large. But, the GT-DPD method outlined above does not need to form the CAM at all nor does it need to do computationally expensive computations of the largest eigenvalue each time. Thus, in the new method, not only has the costly eigenvalue computation been removed, but also the process is distributed among all receiving sensors. When implementing DPD in a scenario where each sensor has limited computational abilities it is more desirable to distribute the load so that no single sensor is over-burdened than to
reduce the total computational complexity across the $L$ sensors. In the original DPD method, the common sensor needs to compute the CAM for each grid point, but since CAM is a Hermitian $L \times L$ matrix formed by CAFs, it just needs to find all the entries on and above the main diagonal. This is equivalent to computing $\frac{L}{2}(L-1)$ CAFs (as non-diagonal elements) and $L$ signal energies (as diagonal elements). Moreover, the common sensor needs to calculate the largest eigenvalue of the matrix for each grid point. On the other hand, in the GT-DPD method, each sensor only needs to compute $(L-1)$ CAFs and one signal energy. Thus, rather than having one sensor compute $\frac{L}{2}(L-1)$ CAFs as in the original DPD, in the method proposed here each sensor computes only $(L-1)$ CAFs; furthermore, each sensor performs a simple Gershgorin estimation rather than a more complicated eigenvalue computation (the number of CAF computations each sensor does is $2/L$ of the amount that the commons sensor does in standard DPD).

The GT-DPD development provides insight into two other previously published alternatives to the classical approach [5], [15]. In the GT-DPD method, if we ignore the maximum operator term in (2.12) and just take $\hat{\lambda}_{\text{max}} = \text{CAF}_i$ for only one arbitrary sensor $i$, then the results will be equivalent to the so-called CAF-Map method, an ad-hoc proposal first presented in [5]. Thus, our Gershgorin-based development provides a sound theoretical rationale for the previously ad hoc development of the CAF-Map method in [5], and also shows that CAF-Map will provide lower accuracy than the full DPD method and the GT-DPD method developed here. Another related approach, called the pair-wise maximum CAF detector, was developed in [15]. It is based on comparing the value of

$$\max_j \max_{\tau,\omega} |\text{CAF}_j(\tau,\omega)|$$

with a threshold $\gamma_{\text{CAF}}$ for only one arbitrary $i$ as reference sensor.
The results in [15] showed that this method also has much lower quality in detection compared to the GLRT detector based on largest eigenvalue; no results were provided in [15] on the location accuracy of the pair-wise maximum CAF method but given its relationship to the GT-DPD method it is clear its performance will not be as good.

The GT-DPD simulation results for many different cases show that the eigenvalue upper bound is very close to the true largest eigenvalue. However, this approximation lowers the quality of the estimation slightly. Most importantly we see that the GT-DPD method retains DPD’s characteristic of reducing the SNR threshold for accurate location. We examined the effect of the proposed approximation on the estimation accuracy using Monte-Carlo computer simulations (with 500 runs each time). In this simulation, a set of 8 moving sensors and one stationary emitter are placed in a configuration as shown in Figure 2.2. There exists a cross ambiguity function for all possible pairs of sensors. We used a BPSK signal with carrier frequency 1GHz as the transmitted signal. The sampling frequency is 80 kHz, symbol rate is 10 kb/s and the number of samples is equal to 4096. We run this simulation for various SNRs (-30, -25, -20, -15, -10, -5, 0 dB) with 500 runs for each SNR. Figure 2.3 shows the effect of the Gershgorin-based eigenvalue approximation on RMS X-Y location error.
2.3. Cross Ambiguity Function Sharing - DPD (CS-DPD)

In this section we develop another approach to implementing Direct Position Determination (DPD) that distributes the CAF computations among all receivers and then
applies data compression to the computed CAFs to facilitate transmitting them to a common site where the DPD location processing is completed. It should be noted that in this method the only approximation that occurs is due to the use of data compression on the CAFs. Thus, in the absence of data compression the accuracy of this method is the same as in the original DPD and for high enough bit rates on the CAF compression the degradation is negligible. However, as the data rates are reduced a drop in performance occurs. In subsection 2.3.1, we first describe the basic method without data compression, and then in subsection 2.3.2 and 2.3.3 specific methods for compressing the computed CAFs are discussed.

2.3.1. Distributed Computation of CAFs

As mentioned in section 2.1.2, the \((i, j)^{th}\) element of the CAM in at grid point \(p\) is the value of the CAF between the signals received by sensors \(i\) and sensor \(j\) at the TDOA/FDOA point corresponding to the grid point \(p\). Thus, sensor \(i\) and sensor \(j\) can contribute to compute the \((i, j)^{th}\) element of the CAM (as well as the \((j, i)^{th}\) element since the CAM is Hermitian). In other words, each pair of sensors \(i\) and \(j\) can share their received signals to compute the \(CAF_{ij}\) in TDOA/FDOA domain. Then, the computed CAFs can be transmitted to the common site; subsections 2.3.2 and 2.3.3 discuss data compression methods for reducing the amount of data needed to be transferred. The common site receives the CAFs and then maps them into the X-Y plane and then uses the results to form the CAM matrix for each grid point and complete the standard DPD localization.

An issue to be considered is how the signal data is shared within the network to allow computation of the needed CAFs. Of course one possibility is for each sensor to transmit
its data to all other sensors, but it is not required by the method. Alternatively, as described below, each sensor only needs to receive raw signal data from a subset of sensors. Figure 2.4 shows a simple case with $L = 4$ sensors where we see how the sensors share their received signals: the solid arrows show the “one step ahead” sharing while the dashed arrows show “two steps ahead” sharing. The matrix shown represents the CAM (which is Hermitian) and the numbers indicate the indices of the specific CAFs needed. The rectangles next to the sensors show which CAFs are computed at which sensor. The diagonal elements of the CAM are equal to the energy of the received signals at each receiver. Although not shown in Figure 2.4, the computed CAFs and the individual received signal energies need to be transmitted to a common site (which likely would be one of the sensors but need not be). Figure 2.5 shows the signal sharing for the case of $L=5$ sensors, which also uses one-step ahead and two-steps ahead sharing. Note that for $L=4$ sensors only half of the sensors must do two-step-ahead sharing while for $L=5$ sensors all five must do two-step-ahead sharing. For both $L = 6$ and $L = 7$ we need up to “three steps ahead” sharing and for $L = 6$ only half of the sensors need to do the highest step sharing while for $L = 7$ all sensors must do that. In general, if we let $S$ be the maximum “step ahead” needed then for $L$ sensors then we need $S = \lfloor L/2 \rfloor$ and if $L$ is odd then all sensors must do that largest step ahead while if $L$ is even then only half the sensors do. Thus, in CS-DPD all the sensors contribute (nearly) equally in the process of computing the CAFs.
The computed CAFs and the sensor received energies are then transmitted to a common site (possibly one of the sensors) for formation of the CAM and the final location processing. Although not required for the method, it is desirable to reduce the amount of data transmission needed for this CAF-sharing step. The next two subsections develop two different data compression methods that exploit specific CAF properties to compress the shared CAFs.
2.3.2. Exploiting CAF Spectral Properties for Data Compression

The CAF is a two-dimensional complex valued function. Thus, we can consider the CAF to be an image (albeit a complex valued image) and apply image compression methods to it. In [6], we used Embedded Zerotree Wavelet (EZW) [27] to independently compress the real part and imaginary part of CAF. Although other wavelet-based data compression methods, like JPEG2000, are also applicable here we focus on EZW because it provides the essential aspects of modern wavelet-based compression without the excessive complexity and flexibility of something like JPEG2000.

Compression causes two kinds of distortion on the CAF data. The first is due to the effect of quantization, which can be modeled by an additive noise. The second is the result of the compression throwing away insignificant coefficients of the wavelet transform, which can be roughly modeled as passing the data through a lowpass filter. Note that a typical CAF contains a large main lobe and (usually) various small side lobes. An important aspect for compression of the CAF is that it is a relatively slowly changing function. Much of the fast changing nature comes from the effect of the additive noise in the received signals. Thus, the important part of the CAF to be retained is a spatially low-pass type “image” that should show up in the medium and low frequency parts of the wavelet transform used in EZW. This is particularly good for the EZW method because it will give raise to many zero tree roots in the EZW processing (see [27] for discussion of zero tree roots). Furthermore, because the any discarded high frequency will contain mostly noise, the EZW algorithm will also tend to perform a de-noising operation to the CAF.
In addition to these general exploitable “image” characteristics of the CAF, there are some other special properties of the CAF that can be exploited for compression. One of the most important properties of CAF is the symmetry. In [8], we have proved that the noise-free CAF is symmetric around its magnitude peak:

\[ |\text{CAF}_y(\tau + \tau_p, \omega + \omega_p)| = |\text{CAF}_y(\tau - \tau_p, \omega + \omega_p)| \]  

(2.13)

where \((\tau_p, \omega_p)\) is the peak of the noise-free CAF magnitude. This symmetry can be exploited for data compression. In practice, the noise slightly perturbs this symmetry so we can rewrite (2.13) as

\[ |\text{CAF}_y(\tau + \tau_p, \omega - \omega_p)| = |\text{CAF}_y(\tau - \tau_p, \omega + \omega_p)| + E \]  

(2.14)

where \(E\) is an error term that is usually small. Thus, using the symmetry property, it is possible to extract the entire CAF magnitude by transmission of only half of the CAF magnitude plus the small residual signal \(E\).

We used Monte Carlo simulations (with 500 runs each time) to examine the performance of the proposed CS-DPD method with EZW compression and also with EZW and symmetry. (Note that without compression the CS-DPD method performs exactly like the original DPD method.) In this simulation, a set of 4 moving sensors and one stationary emitter are placed in a configuration as shown in Figure 2.6. The signals are BPSK with carrier frequency = 1GHz, the sampling frequency = 400 kHz, \(\text{SNR} = -10\) dB and the number of signal samples was 65536. The signals were shared as shown in 2.4 and the six CAFs shown there were computed, compressed and transmitted to a separate
common site (i.e., all CAFs were compressed). Two different compression scenarios were examined. The first scenario applied only the EZW algorithm to compress the CAF (labeled “simple compression” in the figures). The second scenario used the EZW algorithm to compress the CAF together with the symmetry property to reduce the amount of transmitted data (labeled “symmetric compression” in the figures). In this case, the EZW algorithm has been applied only on half of the CAF along with the small residual $E$ (as we see in (2.14)).

![Figure 2.6: Placement of the sensors and the emitter position used for simulation of CS-DPD with EZW-based compression.](image)

The effect of data compression on the RMS error of the location estimation using CS-DPD with compression is shown in Figure 2.7. In this figure, the RMS error is in meter and the data rate is in bits/element, where an “element” is defined as a real (or imaginary) value of a CAF point. As expected the level of error decreases as the bit rate increases,
although the curve is quite flat once above a certain bit rate threshold. Comparing the two curves in each plot shows that the symmetric compression method provides better location accuracy for the same bit rate.
2.3.3. Exploiting CAF Rank Properties for Data Compression

In the previous sub-section we treated the CAF as a (complex-valued) image and applied a compression method tailored for images. Alternatively, the CAF can be viewed as a matrix, which motivates using the singular value decomposition (SVD) for compression. For a complex-valued $M \times N$ matrix $X$, the SVD will be
\[ X = \sum_{i=1}^{r} \sigma_i u_i v_i^H \] (2.14)

where the \( u_i \) are the left singular vectors, the \( v_i \) are the right singular vectors, the \( \sigma_i \) are the nonnegative real singular values ordered such that \( \sigma_i \geq \sigma_{i+1} \), and \( r \) is the number of non-zero singular values. By truncating the above summation to \( k < r \) terms, we get a rank-\( k \) matrix \( X_k \) that approximates \( X \) better than any other rank-\( k \) matrix in the least-square error sense [9], [10]. Thus, from a data compression viewpoint we can think of \( X_k \) as a minimum MSE distortion version of \( X \).

The complex valued \( M \times N \) matrix \( X \) contains \( MN \) complex values or equivalently \( 2MN \) real values. However, the truncated SVD representation \( X_k \) can be represented with only \( k(2M + 2N + 1) \) real values: we need \( kM \) complex values for the \( \{u_i\}_{i=1}^{k} \), \( kN \) complex values for the \( \{v_i\}_{i=1}^{k} \), and \( k \) real values to represent the singular values \( \{\sigma_i\}_{i=1}^{k} \). Thus, assuming the same number of bits are used to represent the elements in each case, the compression ratio achieved in using \( X_k \) rather than \( X \) is:

\[
CR = \frac{MN}{k(M + N + 1/2)}
\] (2.15)

For example, the compression ratio for a \( 128 \times 32 \) matrix truncated with \( k = 1 \) is 25:1.

The CAF usually contains a big main lobe and several small side lobes so that slices at different points we give curves with similar shapes, and this motivates an expectation that a truncation with small \( k \) can give an accurate approximation. This idea is supported by results in [28] where it has been shown that time-frequency operators (which include the CAF) have several large singular values followed by a sharp plunge to smaller values.
and finally decaying asymptotically to zero. Our simulation results also support that most of the information is concentrated in the first few singular vectors/values of the CAF.

The effect of noisy signals is to cause small noise-free singular values to be larger even though most of the true CAF information lies concentrated in the first few singular vectors and values. Thus, SVD truncation has the potential to reduce the amount of noise and equivalently increase the effective SNR. To illustrate this, the singular values of a sample 128x32 CAF are illustrated in Figure 2.8 for two cases: (a) noise-free signals and (b) noisy signals. For the noise-free case, it is clear that there are only 3 to 5 significant singular values, showing that the CAF is very close to a low rank matrix. But for the noisy case the number of significant singular values increases to about 11 or 12. If we truncate the noisy case to have only 3 to 5 terms it is clear that the signal to noise ratio can increase by applying SVD data compression and retaining the first few singular values and discarding the rest.

Figure 2.8: Singular Values of CAF for two cases:(a)Noise-free signal,(b)Noisy signal
We examined the performance of the CS-DPD method using Monte Carlo computer simulations (with 500 runs each time). In this simulation, a set of 4 moving sensors and one stationary emitter are placed in a configuration as shown in Figure 2.9. In this simulation, the signals are BPSK, the sampling frequency = 20 kHz and the number of signal samples was 4096. As in the EZW case, the signals were shared as shown in Figure 2.4 and the six CAFs shown there were computed, compressed and transmitted to a separate common site (i.e., all CAFs were compressed).

Figure 2.10 shows the effect of data compression on RMS error of emitter location estimation for the X and Y dimensions. The four curves compare the cases (i) without compression, (ii) SVD-based compression with compression ratio of 25:1, (iii) SVD-based compression with compression ratio of 8:1, and (iv) SVD-based compression with compression ratio of 5:1. The results show that even for a high compression ratio of 25:1 the estimation accuracy is quite close to the case without compression. Surprisingly, at low SNR the case with the compression ratio of 5:1 (and even the case with the compression ratio of 8:1 in some points) yields more accurate results than without compression case. This improvement is obtained because of the de-noising property of SVD-based data compression.
Figure 2.9: Placement of the sensors and the emitter position used for simulation of CS-DPD with SVD-based compression.

Figure 2.10: Simulation results showing RMS error for X and Y when we compress CAF using SVD-based compression. The four curves compare the cases (i) without compression, (ii) SVD-based compression with compression ratio of 25:1, (iii) SVD-based compression with co compression ratio of 8:1, and (iv) SVD-based compression with compression ratio of 5:1.
2.4. HYBRID CLASSICAL AND DPD (HC-DPD)

In the classical method of TDOA/FDOA localization the CAF computation is typically distributed (and is usually limited to a subset of the total possible CAFs that can be computed) but unlike the CS-DPD method it only transfers the TDOA/FDOA values extracted from the CAFs to the center site rather than complete CAFs as in CS-DPD. Because of the CAF transferal, the DPD method is known to outperform the classical method [14]. In this section, we develop a new method taking the advantages of both classical and DPD scenarios and exploits the relationship between different CAFs to reduce the amount of data transmission in the sensor network.

2.4.1. Relationships Between CAFs from Various Pairs

Suppose that two sensors $S_1$ and $S_2$ receive the ELP signals $\tilde{s}_1(t)$ and $\tilde{s}_2(t)$, respectively. Under the so-called narrowband approximation the ELP model of the received signal for the noise-free case will be [7]:

$$
\tilde{s}_1(t) = \alpha_1 e^{-j\omega_\tau_1} e^{j\omega_t} \tilde{s}(t - \tau_1)
$$

$$
\tilde{s}_2(t) = \alpha_2 e^{-j\omega_\tau_2} e^{j\omega_t} \tilde{s}(t - \tau_2)
$$

(2.16)

where $\tau_i$ is the time delay, $\omega_i$ is the Doppler shift, $\alpha_i$ is the positive-real amplitude and $\omega_c$ is the carrier frequency in rad/sec [7]. Now, we can write one of the received signals in terms of the other one:
\[
\hat{s}_2(t) = \frac{\alpha_2}{\alpha_1} e^{j\omega t} e^{-j\omega t} \hat{s}_1(t + \tau_{12})
\]
\[
\hat{s}_2(t) = G_{12} e^{-j\omega t} \hat{s}_1(t + \tau_{12})
\]
\[
G_{12} = \frac{\alpha_2}{\alpha_1} e^{j\omega_1 t} e^{-j\omega_{12}}
\]

where \( \tau_{12} = (\tau_1 - \tau_2) \) is the TDOA and \( \omega_{12} = (\omega_1 - \omega_2) \) is the FDOA. The CAF between these two signals can be written as being proportional to the Auto Ambiguity function of the first signal as:

\[
CAF_{12}(\tau, \omega) = \int_{-\infty}^{\infty} \hat{s}_1(t) \hat{s}_2^*(t - \tau) e^{-j\omega t} dt
\]
\[
= G_{12}^* e^{-j\omega_2 t} \int_{-\infty}^{\infty} \hat{s}_1(t) \hat{s}_1^*(t + \tau_{12} - \tau) e^{-j(\omega - \omega_{12}) \tau} dt
\]

which gives

\[
CAF_{12}(\tau, \omega) = \gamma_{12}^*(\tau) A_1(\tau - \tau_{12}, \omega - \omega_{12})
\]
\[
\gamma_{12}(\tau) \triangleq G_{12} e^{j\omega_2 \tau}
\] (2.18)

where \( A_1(\tau, \omega) = \int_{-\infty}^{\infty} \hat{s}_1(t) \hat{s}_1^*(t - \tau) e^{-j\omega t} dt \) is Auto Ambiguity function (AAF).

If there is a third received signal, we can similarly show that,

\[
CAF_{23}(\tau, \omega) = \int_{-\infty}^{\infty} \hat{s}_2(t) \hat{s}_3^*(t - \tau) e^{-j\omega t} dt
\]
\[
= \int_{-\infty}^{\infty} G_{12} e^{-j\omega_2 t} \hat{s}_1(t + \tau_{13}) \hat{s}_3^*(t - \tau) e^{-j\omega \tau} dt
\]
\[
= \gamma_{12}(\tau_{12}) e^{j\omega_2 \tau} CAF_{13}(\tau + \tau_{12}, \omega + \omega_{12})
\] (2.19)

Further, by putting \( \tau = \tau_{12} \) and \( \omega = \omega_{12} \) in (2.18), we can compute \( \gamma_{12}(\tau_{12}) \) as,
\[
\gamma_{12}(\tau_{12}) = \frac{\text{CAF}_{12}^*(\tau_{12}, \omega_{12})}{A_{11}^*(0,0)}
\]

(2.20)

In general, we can show that:

\[
\text{CAF}_{\text{in}}(\tau, \omega) = \gamma_{\text{in}}(\tau_{\text{in}}) e^{i\omega_{\text{in}}} \text{CAF}_{\text{in}}(\tau + \tau_{\text{in}}, \omega + \omega_{\text{in}})
\]

(2.21)

Similarly, we can use \( \gamma_{\text{in}}(\tau_{\text{in}}) = \frac{\text{CAF}_{\text{in}}(\tau_{\text{in}}, \omega_{\text{in}})}{A_{11}(0,0)} \) to compute the term \( \gamma_{\text{in}}(\tau_{\text{in}}) \) in (2.21).

2.4.2. Using CAF Relationships for Hybrid Classical - DPD

Exploiting the (2.21), it is possible to approximately construct the CAM by transferring only its first row plus the corresponding estimated TDOA/FDOAs (estimated as they are in the classical method). Thus, this distributed method is a hybrid combination of the classical method and the DPD method so we call it the Hybrid Classical – DPD method or HC-DPD.

Figure 2.11 illustrates a simple case of 4 receiving sensors. As we see in Figure 2.11-(a), sensor 1 (any of the sensors could play this role, we chose sensor 1) distributes its received signal to the other sensors in the network, each of which then computes the CAF between its own signal and the signal from sensor 1. Then, each sensor can transmit the computed CAF to a common site (which can be one of the sensors). In Figure 2.11-(b), we assumed sensor 2 as the common point. After receiving the CAFs, sensor 2 is able to compute \( \text{CAF}_{23} \), having \( \text{CAF}_{12} \) (computed by itself) and \( \text{CAF}_{13} \) (received from sensor 3) applying (2.19) and (2.20). In (2.19), \( \tau_{12} \) and \( \omega_{12} \) are estimated by peak of the \( \text{CAF}_{12} \) magnitude and \( \gamma_{12}(\tau_{12}) = \frac{\text{CAF}_{12}^*(\tau_{12}, \omega_{12})}{A_{11}^*(0,0)} \) where \( A_{11}(0,0) \) is the value of the auto
ambiguity function of the received signal by sensor1 at its magnitude peak (that can be also computed by sensor2 or can be transferred to sensor2 as a number).

Similarly, sensor 2 is also able to compute $CAF_{24}$ and $CAF_{34}$ having $CAF_{12}$, $CAF_{13}$ and $CAF_{14}$ applying (2.21). Now, sensor 2 has everything needed to approximately form the CAM and do the location estimation.

![Diagram](image)

Figure 2.11: HC-DPD for four receivers.

We examined the proposed method using Monte-Carlo computer simulations (with 500 runs each time). In this simulation, a set of 3 moving sensors and one stationary emitter were placed in a configuration as shown in Figure 2.12. In this simulation we computed $CAF_{12}$ and $CAF_{13}$ directly using Stein’s method [1] and then computed the $CAF_{23}$ using (2.19). Then having all of the elements, we formed matrix CAM and apply the DPD method. The signals are BPSK with carrier frequency 1GHz, the sampling frequency is 20 kHz and the number of samples is equal to 4096. Figure 2.13 shows the RMS error of emitter location estimation for X and Y dimensions compared to original DPD and Classical methods.
HC-DPD leads to a large reduction in computation load compared to DPD and large reduction in the amount of data transmission compared to CS-DPD but with the cost of lower accuracy; however, as we can see in Figure 2.13, it is still much more accurate compared to classical TDOA/FDOA method.

Figure 2.12: Placement of the sensors and the emitter position used for simulation.

Figure 2.13: RMS errors for X and Y estimates versus SNR
2.5. **Comparison between DPD and the Proposed Methods**

The proposed methods are summarized in the following paragraphs as well as in Table 2.1, where “+” indicates advantages and “-“ indicates disadvantages.

The **GT-DPD method** exploits the simplicity of the Gershgorin theorem to approximately compute the largest eigenvalue without the high cost of exactly computing it. This enables each sensor to locally make its best estimate of the location based on that data it has. These locally-generated estimates are then transmitted to a central location where a final decision is made. Use of the Gershgorin Theorem theorem enables nearly the same accuracy as the original DPD method while being exceptionally good at distributing the computations; data transmission requirements, however, are on par with DPD.

In **CS-DPD method**, we divide the computational load among all sensors of the network. In this method, we don’t use any other approximation more than data compression and the quality of this method is the same as the Original DPD for proper bit rates. The extra transmission required is mitigated by data compression so the final transmission requirements are on par with DPD.

Finally, in the **HC-DPD Method**, we used the same idea of Decentralized DPD in addition to exploiting the relationship between different CAFs to eliminate the data redundancy and this idea leads to a large data transmission reduction but at the cost of a reduction in accuracy.
Table 2.1: Comparison between DPD and the Proposed Methods

<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy</th>
<th>Distributed Computation</th>
<th>Data Transmission</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPD</td>
<td>++</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>GT-DPD</td>
<td>++</td>
<td>++</td>
<td>-</td>
</tr>
<tr>
<td>CS-DPD</td>
<td>++</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>HC-DPD</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
Chapter 3

A Novel Method: Spatial Sparsity Based Localization in Wireless Sensor Network

3.1. Localization Based on Spatial Sparsity, and Comparison with Classic and DPD methods

As mentioned before, the classic two-stage localization method is not necessarily optimal because in the first stage of these methods, the TDOA and FDOA estimates are obtained by ignoring the fact that all measurements should be consistent with a single emitter location. In other words, each stage can be itself optimal but the cascade of the two stages is not necessarily optimal.

We have developed a novel one-stage TDOA/FDOA localization method based on spatial sparsity of emitters in x-y plane (or x-y-z space in 3-dimensional localizations) [20], [21]. In this method, we directly estimate the location of the emitter(s) without going through the intermediate stage of TDOA or FDOA estimation. It is obvious that in emitter location problems, the number of emitters is much smaller than the number of grid points in a fine grid on the x-y plane. Thus, if we assign a positive number to the grid cells containing an emitter and assign zero to the rest of grid cells; this yields a very
sparse grid plane matrix that can be reformed as a sparse vector. Since each element of this vector corresponds to one grid point in the x-y plane, we can estimate the location of emitters by extracting the position of non-zero elements of the sparest vector that minimize the cost between true received signals and expected received signals that satisfy the TDOA/FDOA relationships.

In [29], the authors suggested a source localization method based on TDOA in a multipath channel exploiting the sparsity of the multipath channel for estimation of the line-of-sight component. In this method, the sensors don’t need to know the information on the specific transmitted symbols but, they require knowledge of the pulse shape of the transmitted signal. In [30], the authors suggested a compressive-sensing based target localization using TDOA. In this method, each sensor approximates the transmitted signal by its own received signal mapped to each one of the grid points.

In this research, contrary to [29] and [30], we developed a method based on both TDOA and FDOA to take advantage of both delay and Doppler shifts. Contrary to [29], our method does not need any knowledge of the transmitted signal’s pulse shape nor any other a priori information, and unlike [30], we consider the transmitted signal as a deterministic unknown signal that will be estimated in the sensor network using all received signals.

The proposed method is evaluated in different scenarios including localization in the presence of multipath, simultaneous localization for multiple emitters and also localization for multiple emitters in the presence of multipath. Unlike the single-step DPD method in [12], [13], [14], [15], our method handles multiple emitters and multipath scenarios quite well.
3.2. Sparsity Approach

In principle, sparsity of the grid vector can be enforced by minimizing its $\ell_0$-norm (i.e., the number of non-zero elements in the grid vector). However, since the $\ell_0$-norm minimization is an NP-hard non-convex optimization problem, it is very common (e.g. in sparse problems and related compressive sensing problems) to approximate it with $\ell_1$-norm minimization, which is a convex optimization problem and also achieves the sparse solution very well [31]. The typical $\ell_1$-norm minimization is defined as,

$$\begin{align*}
\hat{s} &= \arg\min \|s\|_1 \\
st &\quad \Theta \times s = r 
\end{align*}$$

(3.1)

where $\|v\|_p = \sqrt[p]{\sum |v|^p}$.

Note that unlike $\ell_1$-norm minimization, $\ell_2$-norm minimization which minimizes the signal energy, almost never achieves a sparse solution [31]. Baraniuk used a simple geometry model that helps visualize why $\ell_2$-norm minimization fails to achieve the sparse solution [31]. Figure 3.1-(a) demonstrates the $\ell_2$-norm minimization problem. In this figure, the green hyper-plane H represents the constraint (translated null space of $\Theta$). The $\ell_2$-norm minimization aims to find the point on H having the minimum distance to the origin. To this end, we can imagine blowing up a hyper-sphere until it touches the H at $\hat{s}$. As we see in the figure, the closest point $\hat{s}$ will be away from the coordinate axes with high probability, and thus will not be sparse. On the other hand, Figure 3.1-(b) demonstrates the $\ell_1$-norm minimization. Since $\ell_1$-ball has points aligned with the
coordinate axes, if we blow up the $\ell_1$-ball, it will first touch the hyper-plane $H$ at a point on the coordinate axes, which provides the sparse solution [31].

![Image of coordinate axes and hyper-plane](image)

Figure 3.1: $\ell_2$-norm and $\ell_1$-norm minimization visualization [31].

Thus, after formulating the localization problem in terms of the sparse grid vector, we can estimate this vector by pushing sparsity using $\ell_1$-norm minimization on the grid vector, and at the same time minimizing the cost between the actual received signals be sensors and the predicted received signal estimations that satisfy the TDOA/FDOA relationships.

### 3.3. Problem Formulation

Suppose that an emitter transmits a signal and $L$ sensors receive that signal. The complex baseband signal observed by the $l$th sensor is

$$r_l(t) = \alpha_l e^{j\phi_l} s(t - \tau_l) e^{j2\pi f_l t} + w_l(t)$$  \hspace{1cm} (3.2)
where \( s(t) \) is the transmitted signal, \( \alpha_l \) is the real path attenuation, \( \varphi_l = -2\pi f_c \tau_i \) is constant phase, \( f_c \) is the carrier frequency, \( \tau_i \) is the signal delay, \( f_l \) is the Doppler shift, and \( w_i(t) \) is a white, zero mean, complex Gaussian noise \([7]\). Assume that each sensor collects \( N_s \) signal samples at sampling frequency \( F_s = 1/T_s \). Then, we have

\[
\mathbf{r}_i = \alpha_l e^{j\omega_l} \mathbf{W}_l \mathbf{D}_l \mathbf{s} + \mathbf{w}_i
\]

(3.3)

where \( \mathbf{r}_i \) is the vector containing \( N_s \) samples of the received signal by \( l \)th sensor, \( \mathbf{s} \) is \( N_s \) samples of the transmitted signal, \( f_l \) is the Doppler shift and \( \mathbf{D}_l \) is the time sample shift operator by \( n_l = (\tau_i / T_s) \) samples. We can write \( \mathbf{D}_l = \mathbf{D}^{n_l} \) where \( \mathbf{D} \) is an \( N_s \times N_s \) permutation matrix defined as \([\mathbf{D}]_{ij} = 1 \) if \( i = j + 1 \), \([\mathbf{D}]_{0, N-1} = 1 \) and \([\mathbf{D}]_{0j} = 0 \) otherwise.

\[
\mathbf{D} = \begin{bmatrix}
0 & 1 & 0 & \ldots & \ldots & 0 \\
1 & 0 & 1 & \ldots & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 1 & 0 & 1 & 0 & 0 \\
\end{bmatrix}, \quad \mathbf{D}_l = \begin{bmatrix}
0 & 1 & 0 & \ldots & \ldots & 0 \\
1 & 0 & 1 & \ldots & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 1 & 0 & 1 & 0 & 0 \\
\end{bmatrix}^{n_l}
\]

(3.4)

Now, let \( z_{x,y} \) be a number corresponding to the grid point at \((x,y)\), where the value is one if the \((x,y)\) grid point contains an emitter and is zero if it does not. Thus, the signal vector received by \( l \)th sensor will be

\[
\mathbf{r}_l = \sum_{x} \sum_{y} z_{x,y} \alpha_{l,x,y} e^{j\varphi_{l,x,y}} \mathbf{W}_{l,x,y} \mathbf{D}_{l,x,y} \mathbf{s} + \mathbf{w}_l
\]

(3.5)
where $\alpha_{l,x,y}$, $\varphi_{l,x,y}$, $W_{l,x,y}$ and $D_{l,x,y}$ are the real path attenuation, constant phase, Doppler shift operator and time sample shift operator respectively, with respect to sensor $l$ assuming that the emitter is located in the grid point $(x,y)$. The summations are over all grid points in the desired $(x,y)$ range. Now, if we reform all of the grid points in a column vector and re-arrange the indices, we will have

$$r_l = \sum_{n=1}^{N} z_n \alpha_{l,n} e^{j\varphi_{l,n}} W_{l,n} D_{l,n} s + w_l. \quad (3.6)$$

In (3.6), we consider the transmitted signal $s$ as a deterministic unknown signal (a common signal model in localization problems). Then, for each grid point, we estimate the transmitted signal using the Minimum Variance Unbiased estimator (MVU) as

$$\hat{s}_n = \frac{1}{L} \sum_{l=1}^{L} e^{-j\varphi_{l,n}} D_{l,n}^{-1} W_{l,n}^{-1} r_l, \quad (3.7)$$

where $\hat{s}_n$ is the MVU estimate for the transmitted signal from grid point $n$. Note that the average effect of path attenuations are embedded in the signal estimate $\hat{s}_n$.

We define the matrix $\Gamma_n$ as the Doppler and delay operator with respect to all $L$ sensors, assuming that the received signal comes from the grid point $n$ (there is an emitter at grid point $n$):

$$\Gamma_n \triangleq \left[ e^{j\varphi_{1,n}} W_{1,n} D_{1,n} e^{j\varphi_{2,n}} W_{2,n} D_{2,n} \cdots \right]_{L \times N_s} \quad (3.8)$$
Then, we can define \( \theta_n, n \in \{1, 2, \ldots, N\} \) as an \( LN_s \times 1 \) vector containing all signals received by all \( L \) sensors when the emitter is in grid point \( n \) as

\[
\theta_n \triangleq \Gamma_n \times \hat{s}_n.
\] (3.9)

If we arrange all vectors \( \theta_n \) for \( n: 1 \ldots N \) as the columns of a matrix \( \Theta \) as

\[
\Theta = [\theta_1 \ \theta_2 \ \ldots \ \theta_N]_{LN_s \times N},
\] (3.10)

and then, we have

\[
r = \Theta \times z + w
\] (3.11)

\[
r \triangleq [r_1^T \ r_2^T \ \ldots \ r_L^T]^T_{LN_s \times 1}
\]

\[
z \triangleq [z_1 \ z_2 \ \ldots \ z_N]^T_{N \times 1},
\]

where \( r \) is the vector of all \( L \) received signals, \( z \) is the sparse vector of \( z \)-values assigned to each grid point and \( w \) is the noise. Now, we can solve our problem by forming a BPIC (Basis Pursuit with Inequality Constraints) problem [32] as following:

\[
\begin{cases}
\hat{z} = \arg \min \|z\|_p \\
\text{s.t} \quad \|\Theta \times z - r\|_2 \leq \varepsilon
\end{cases}
\] (3.12)

or regularized BPDN (Basis Pursuit Denoising) problem [32] as:

\[
\hat{z} = \arg \min \|\Theta \times z - r\|_2 + \lambda \|z\|_p
\] (3.13)

where \( \| \cdot \|_p \) is the \( \ell_p \)-norm defined as \( \|v\|_p = \sqrt[|p|]{\sum |v|^p} \) and \( \lambda \) is the regularization parameter that adjust the balance between cost (fit) and sparsity. Selecting a proper value for \( \lambda \) is important in regularization problems; by increasing \( \lambda \), the penalty on cost gets higher and
it leads to a good fit to the observation. However, if we set \( \lambda \) too high, it may fail to achieve a sparse solution. Various methods have been proposed in the literature for automated \( \lambda \) selection (such as in [33]). However, it is usually possible to find reasonable choices for \( \lambda \) by empirical methods.

3.4. Comparing Spatial Sparsity Based Method with Classic and DPD methods

We examined the performance of the proposed method and compared the results using Monte-Carlo computer simulations for different scenarios. In the first simulation, we assumed that 3 moving sensors receive the signal from one stationary emitter placed in a configuration as shown in Figure 3.2 (the location of the emitter has been chosen randomly). In this simulation, we used a BPSK signal with carrier frequency of 1 GHz as the transmitted signal. The sampling frequency is 80 kHz, symbol rate is 10 kb/s and the number of samples is equal to 1024. In Figure 3.3, we can see the RMS Error vs. SNR (with 500 runs for each SNR) for estimating the location of the emitter in \((x−y)\) plane. As we see, the proposed method has better performance compared to DPD and Classic methods.
Figure 3.2: Placement of the sensors and the emitter position used for simulation.
One of the challenging topics in emitter location problems is robustness in the presence of multipath reflections. We evaluated the capability of the proposed method in dealing with multipath scenarios using Monte-Carlo simulation. In this simulation, we
assumed that 4 moving sensors receive the signal from one stationary emitter placed in a configuration as shown in Figure 3.4 (the location of the emitter has been chosen randomly). The position of the reflectors are arbitrarily set at (2703,4470), (4084,4674), (4203,3928), (3160,2624). Figure 3.4 shows the RMS Error vs. SNR (with 500 runs for each SNR) for estimating the location of the emitter in multipath case. The following plots show better accuracy of the proposed method over DPD and Classic methods. As we see, none of the DPD and Classic methods provide unbiased estimates for higher SNRs.

Figure 3.4: Placement of the sensors and the emitter position used for simulation.
In another simulation, we evaluated the performance of the proposed method and compared the results to DPD and Classic method for multi-emitter scenarios when we have more than one transmitter in the range of interest and we aim to estimate the location of all emitters. In this simulation, we assumed that 6 moving sensors receive the
signal from two stationary emitters placed in the configuration as shown in Figure 3.6 (the location of the emitters has been chosen randomly). Note that in the proposed method, we don’t need to know the number of emitters because we achieve a clear sharp spike for each emitter in estimated sparse vector. Thus, the number of the spikes corresponds to the number of emitters. Figure 3.7 shows the RMS Error vs. SNR (with 500 runs for each SNR) for estimating the location of two emitters. Comparing the RMS Error curves, we see that the proposed method achieves much better accuracy with significantly smaller error compared to DPD and Classic methods. As we see, similar to multipath case, both DPD and Classic methods provide biased estimates at high SNRs.

![Figure 3.6: Placement of the sensors and the emitter position used for simulation.](image)
Figure 3.7: RMS Error for X and Y (meter) versus SNR (dB) for estimating the locations of two emitters in (x-y) plane (multi-emitter scenario). (a), (b): RMSE for first emitter. (c), (d): RMSE for second emitter.

Finally, we evaluated the performance of the proposed method for multiple emitter localization in the presence of multipath. In this simulation, we assumed that 8 moving sensors receive the signal from two stationary emitters placed in the configuration as shown in Figure 3.8 (the location of the emitters has been chosen randomly). In this simulation, we used a signal with carrier frequency 100 MHz as the transmitted signal. The sampling frequency is 1 MHz, symbol rate is 30 kb/s and the number of samples is equal to 1024. Figure 3.9 shows the RMS Error vs. SNR (with 500 runs for each SNR)
for estimating the location of two emitters. Again, we can see the higher performance of the proposed method with significantly smaller error compared to DPD and Classic methods.

Figure 3.8: Placement of the sensors and the emitter position used for simulation.
Figure 3.9: RMS Error for X and Y (meter) versus SNR (dB) for estimating the locations of two emitters in (x-y) plane (multi-emitter scenario) in the presence of multipath. (a), (b): RMSE for first emitter. (c), (d): RMSE for second emitter.

In the proposed method, since the emitter location is estimated based on the original received signals altogether, rather than the parameters derived individually from each signal, the proposed method shows better performance compared to two-stage methods. Furthermore, the effect of multipath will be degraded in this method, because the indirect signal reflections are usually not consistent with a single location. Thus, the multipath signals combine incoherently in the signal estimation. Pushing the sparsity in position domain also helps to reduce multipath effect, and make the algorithm capable to
distinguish between locations of various emitters in multiple-emitter scenarios. The DPD method optimizes only with respect to the $\ell_2$-norm of the error; the sparsity-based method proposed here exploits the additional $\ell_1$-norm to force better accuracy – especially in the presence of multipath and multiple emitters.

As we see, the simulations for different scenarios including single-emitter and multiple-emitters cases with or without multipath reflections show the high performance and accuracy of the proposed method (especially in multipath conditions and multiple-emitters cases) compared to other methods. Simulation results show that contrary to DPD and Classic methods, the proposed method is a very reliable and strong tool to deal with multipath and multiple emitter scenarios.
Chapter 4

Indoor Localization, Tracking and Fall Detection for Assistive Healthcare Based on Spatial Sparsity

4.1. Tracking of Patients and Fall Detection for Assistive Healthcare

Indoor localization has been a long-standing and important issue in the areas of signal processing and sensor networks that has raised increasing attention recently [34] - [43]. As the number of elderly people grows rather quickly over the past few decades and continues to do so [44], it is imperative to seek alternative and innovative ways to provide affordable health care to the aging population [45]. A compelling solution is to enable pervasive healthcare for the elderly and people with disabilities in their own homes, while reducing the use and dependency of healthcare facilities. To this aim new technology and infrastructure must be developed for an in-home assistive living environment. One of the key demands in such an assistive environment is to promptly and accurately determine the state and activities of an inhabitant subject. The three-dimensional indoor localization provides an effective means in tracking the position, motions, reactions and fall detection
of a patient, the elderly or any person with special needs for medical observation or accident prevention.

In assistive healthcare applications, the individual may wear a small device that could emit a radio frequency (RF) signal for localization. This emitter(s) propagates a signal that could be received and captured by several pre-mounted wireless sensors located in known positions. The sensors can estimate the location of the emitter after sharing some data and performing some processing.

As mentioned before, the classic approach for localization is to first estimate one or more location-dependent signal parameters such as time-of-arrival (TOA), angle-of-arrival (AOA) or received-signal-strength (RSS). Then in a second step, the collection of estimated parameters is used to determine an estimate of the emitter’s location. However, the systems based on AOA need multiple antennas or a scannable antenna that are usually costly [36]. The methods based on signal strength measurement (RSS) require a costly training procedure and complex matching algorithms, while the positioning accuracy of these methods is also limited by the large variance in indoor environments [37], [43]. The methods based on time-of-arrival (TOA) are usually very accurate. However, the accuracy of the classic TOA based methods often suffers from massive multipath conditions in indoor localization, which is caused by the reflection and diffraction of the RF signal from objects (e.g., interior wall, doors or furniture) in the environment [34].

In this section, we apply the proposed sparsity based method for indoor localization exploiting spatial sparsity of the emitter in the x-y-z space to estimate the location of the emitter directly without going through the intermediate stage of TOA estimation. We will see that this method is very robust and effective in dealing with the multipath condition,
which is a very serious problem in indoor localization due to the many reflections from furniture and walls. In our method, we don’t need any time synchronization between emitter and receivers since the method is implicitly based on time difference of arrival (TDOA) between receivers. In this approach, the height of the worn device (emitter) from the floor is also estimated. Thus, the system can immediately detect if the patient falls on the floor because of unconsciousness or any other reason.

We evaluate the performance of the proposed method using various Monte Carlo computer simulations for different cases. The simulation results show the accurate and fast localization performance of this method even in multipath conditions, with low SNRs and with small number of sensors; this provides a significant advantage over TOA, RSS or other single-stage methods.

4.2. Problem Formulation

Assume that we divide up the x-y-z space into fine enough grids. It is obvious that in emitter location problems, the number of emitters is much smaller than the number of all grid points in the x-y-z space. Thus, by assigning a positive number to each one of the grids containing an emitter and assigning zeros to the rest of the grid cells, we will have a very sparse 3-dimensional grid matrix that can be reformed as a sparse vector. We can estimate the location of emitters by extracting the position of non-zero elements of the sparse vector. To this end, we have to estimate the sparsest vector that minimizes the cost between predicted received signals and actual received signals with respect to the signal model and delay relationship between transmitted signals and received signals. As
mentioned before, the sparsity will be achieved by minimizing the \( \ell_1 \)-norm of the grid vector.

Suppose that an emitter transmits a signal and \( L \) sensors receive that signal. The complex baseband signal observed by the \( l \)th sensor is

\[
r_l(t) = \alpha_l s(t - \tau_l) + w_l(t)
\]  

where \( s(t) \) is the transmitted signal, \( \alpha_l \) is the complex path attenuation, \( \tau_l \) is the signal delay and \( w_l(t) \) is a white, zero mean, complex Gaussian noise. Assume that each sensor collects \( N_s \) signal samples at sampling frequency \( F_s = 1/T_s \). Then we have

\[
\mathbf{r}_l = \alpha_l \mathbf{D}_l \mathbf{s} + \mathbf{w}_l ,
\]

where \( \mathbf{r}_l \) is the vector containing \( N_s \) samples of the received signal by \( l \)th sensor, \( \mathbf{s} \) is \( N_s \) samples of the transmitted signal and \( \mathbf{D}_l \) is the time sample shift operator by \( n_i = (\tau_i / T_s) \) samples. We can write \( \mathbf{D}_l = \mathbf{D}^\nu \) where \( \mathbf{D} \) is an \( N_i \times N_i \) permutation matrix defined as

\[
[\mathbf{D}]_{ij} = 1 \text{ if } i = j + 1 , \quad [\mathbf{D}]_{k,N-1} = 1 \quad \text{and} \quad [\mathbf{D}]_{kj} = 0 \text{ otherwise}.
\]

Now, we assign a number \( \delta_{x,y,z} \) to each one of the grid points \((x,y,z)\). Assume that \( \delta_{x,y,z} \) is one for the grid points containing an emitter and zero for the rest of the grid points. Thus, the signal vector received by \( l \)th sensor will be

\[
\mathbf{r}_l = \sum_x \sum_y \sum_z \delta_{x,y,z} \alpha_{l,x,y,z} \mathbf{D}_{l,x,y,z} \mathbf{s} + \mathbf{w}_l
\]  

(4.3)
where \( D_{l,x,y} \) is the time sample shift operator with respect to sensor \( l \) assuming that the emitter is located in the grid point \((x,y,z)\) and the summations are over all grid points in the desired \((x,y,z)\) range. Now, if we reform all of the grid points in a column vector and re-arrange the indices, we will have

\[
\mathbf{r}_l = \sum_{n=1}^{N} z_n \alpha_{l,n} D_{l,n} \mathbf{s} + \mathbf{w}_l. \tag{4.4}
\]

where \( N \) is the total number of grid points in the \((x,y,z)\) range of interest.

In assistive healthcare systems, we can easily assume that the transmitted signal \( s \) is known by the receiver sensors. However, in other applications when the signal is not known for receivers, we can consider the transmitted signal \( s \) as a deterministic unknown signal. Then, for each grid point, we estimate the transmitted signal using the Minimum Variance Unbiased estimator (MVU) as

\[
\hat{s}_n = \frac{1}{L} \sum_{l=1}^{L} D_{l,n}^{-1} \mathbf{r}_l. \tag{4.4}
\]

where \( \hat{s}_n \) is the MVU estimate for the transmitted signal from grid point \( n \).

We define the matrix \( \Gamma_n \) as the delay operator with respect to all \( L \) sensors, assuming that the received signal comes from the grid point \( n \) (i.e. an emitter exists at grid point \( n \)):

\[
\Gamma_n = \begin{bmatrix}
\alpha_{1,n} D_{1,n} \\
\alpha_{2,n} D_{2,n} \\
\vdots \\
\alpha_{L,n} D_{L,n}
\end{bmatrix}_{L \times N_i}
\]

Then, we can define \( \theta_n, n \in \{1, 2, ..., N\} \) as an \( LN_i \times 1 \) vector containing all signals received by all \( L \) sensors when the emitter is in grid point \( n \) as
\[ \theta_n = \Gamma_n \times \hat{s}_n \]  

(4.6)

If we arrange all vectors \( \theta_n \) for \( n:1...N \) as the columns of a matrix \( \Theta \) as

\[ \Theta = [\theta_1 \theta_2 \ldots \theta_N]_{L N \times N} \]  

(4.7)

then, the received signals are given by,

\[ r = \Theta \times v + w, \]  

(4.8)

\[ r = [r_1^T \ r_2^T \ldots \ r_L^T]_{L N \times 1} \]

\[ v = [\delta_1 \ \delta_2 \ldots \ \delta_N]^T_{N \times 1} \]

where \( r \) is the vector of all \( L \) received signals, \( v \) is the sparse vector of \( \delta \)-values assigned to each grid point, and \( w \) is the noise. Now, we can solve our problem by forming a regularized BPDN (Basis Pursuit Denoising) problem as:

\[ \hat{v} = \arg \min \| \Theta \times v - r \|_2 + \lambda \| v \|_1 \]  

(4.9)

where \( \| \cdot \|_p \) is the \( \ell_p \)-norm defined as \( \| v \|_p = \sqrt[p]{\sum |v|^p} \) and \( \lambda \) is the regularization parameter balancing the sparsity versus estimation cost.

In order to reduce the computational complexity, we can start with a coarse grid, and then increase the resolution of the grid iteratively. In other words, after estimating the approximate location of the target in a coarse grid space, we generate a finer grid in a smaller area of interest around the approximately estimated position, and then re-calculate the location in this new area. We continue this process until we achieve the desired positioning accuracy.
Furthermore, to reduce the computational complexity and achieve more accurate results in real-time patient tracking, we can exploit the prior position information, to estimate the new position of the patient. To this end, we only need to start the localization process over the entire grid once to estimate the initial position. After that, we can always restrict the area of interest to only a small vicinity around the past estimated position, considering the maximum possible movement of the patient. This technique leads to a significant reduction in size of the sparse vector and computational complexity, and also helps to achieve more accurate results by choosing a finer grid in the area of interest.

4.3. Simulation Results for Indoor Localization and Fall Detection

We examined the performance of the proposed method using various computer simulations in different scenarios. In the first scenario, we ran the Monte Carlo simulation with 500 runs each time for various numbers of sensors (from 3 to 8 sensors). We simulated the multipath conditions in a typical apartment shown in Figure 4.1-(a). The sensors are mounted on the floor at $x$-$y$ locations $(0,0)$, $(0,10)$, $(10,0)$, $(10,10)$, $(0,4)$, $(4,10)$, $(10,6)$, $(6,0)$ respectively and the location of the target has been chosen randomly. In this simulation, we used a BPSK signal with carrier frequency of 1 GHz. The sampling frequency is 200 KHz and the number of samples is equal to 256. We run this simulation one time for $\text{SNR} = 0\text{dB}$ and another time for $\text{SNR} = 10\text{dB}$.

Figure 4.1-(b) shows the same apartment with four of the receiver sensors (i.e., red dots) mounted at the corners. Figure 4.1-(c) illustrates a simple example for multipath scenario. In this figure, the solid (blue) lines present the direct paths (line-of-sight) and
the dashed (red) lines indicate the reflected paths. However, given the extremely complex nature of the reflections within such a closed environment and the tremendous difference in the reflection rates for different building materials, it is impossible to conclude a rather perfect multi-path reflection model for the indoor circumstance. However, it is well agreed that the strength of reflected signals deteriorates after each reflection. Moreover, the TOA based localization systems usually suffer from first-order reflections since they generate the side-lobes very close to the main peak in the correlation stage used in traditional TOA based methods. Thus, the models like in Figure 4.1-(c) seem reasonable for the purpose of research.

Figure 4.2 shows the RMS Error vs. number of sensors for estimating the location of the target. As we expected, the accuracy gets better by increasing the number of the receiving sensors. The results demonstrate that the proposed method has very good performance even for small number of sensors (e.g. 3 sensors). This finding enables the possibility of using small number of sensors to reduce the complexity and expenses of the system.
Figure 4.1: (a) A typical apartment layout. (b) Four sensors mounted in the corners. (c) A simple case for multipath scenario. The solid line is direct path and dashed lines are reflected paths.
Figure 4.2: RMS Error for X and Y (meter) versus Number of sensors.
In the second scenario, we aim to track a moving target (e.g. a patient walking in the apartment). Again, we simulated the multipath conditions, and we used only 4 sensors mounted on the corners of a building shown in Figure 4.3-(a) for the purpose of localization. In this simulation, we used a BPSK signal with carrier frequency of 40 MHz. The sampling frequency is 80 KHz, the number of samples is equal to 256, and SNR = 10dB. Figure 4.3-(a) shows the actual path (blue line) of the patient walking around in a building, as well as the estimated path (red line) by the system. Figure 4.3-(b) shows the error in positioning defined as

\[ e = \sqrt{e_x^2 + e_y^2 + e_z^2} \]

(4.11)

where \( e_x, e_y, \text{ and } e_z \) are RMS errors in positioning for X, Y and Z dimensions. Note that the position of the person is not necessarily located on the grid points. However, the algorithm always finds the nearest grid point to the patient’s location. Thus, a significant portion of the errors in Figure 4.3-(b) is resulted from the distance between the target and the determined nearest grid point close to it. However, this kind of error can be reduced by employing finer grids.
Figure 4.3: (a) True position of the patient (in blue) and the estimated position (in red), (b) Error in positioning for each location in part (a).

In the third scenario, we used the proposed localization technique to detect the person’s fall, which is one of the most serious sources of injury for elderly people [46], [47], [48]. As mentioned before, we are always able to detect falls using the proposed 3-dimensional localization technique. In this method, we obtain the accelerations in all
three directions by computing the second derivative of the wearable emitter position versus time, and then compared the norm of the accelerations against the predefined thresholds. For example, the second derivative of the distance from the floor (versus time) can present the acceleration in \( z \)-direction. The acceleration thresholding technique is an effective method for fall detection with a low false-positive rate [47].

Furthermore, the position of the person in \( (x-y-z) \) space can help to reduce the rate of missed detections (false-negatives). For example, we don’t expect the person to lie down on the floor in the kitchen. Thus, by detecting this sort of positions in \( (x-y-z) \) space, the system can interpret that a fall has happened. This is the advantage of this method over traditional acceleration based fall detectors.

Figure 4.4 shows the actual path (blue line) of the patient as well as the estimated path (red line) in 3-dimensional space. As we see, the system can track the exact position of the wearable device in 3-dimensional space, and utilize it for the purpose of fall detection.
Figure 4.4: True position of the patient (in blue), the estimated position (in red), and the detected position of the fall.

As mentioned before, the indoor localization is a very beneficial tool in assistive healthcare environment when tracking the locations, behaviors and reactions of the patient is required for medical observation, symptoms identification or accident prevention. On the other hand, fall detection is an effective approach to minimize the effects of the patient or elderly falls by reporting it and calling for help. The simulation results show that the proposed method has very good performance even with small number of sensors. The results also indicate that, in contrary to the classic TOA based methods, the proposed approach is a very effective and robust tool to deal with multipath issues. Furthermore, the system works very well in noisy environments with low SNRs. It implies that, even with low transmitted power (to keep the worn device small with long battery life), we can still achieve a high localization accuracy.
Chapter 5
In-Body Localization Using Wireless Body Sensor Network: Applications and Challenges

5.1. Importance and Applications

As the main part of this dissertation, we explore innovative technologies for sparse signal processing in biomedical systems including sparsity-based in-body localization methods. In this research, we develop the sparsity based method for accurate in-body implant localization using wireless body sensor network. This method is based on spatial sparsity in 3D space and convex optimization theory to achieve very precise results for implant localization with less than 1 mm error in positioning. The proposed method is more accurate and easier to implement compared to other existing methods, and compliant with the Medical Implant Communication Services (MICS) standards. Furthermore, the system demonstrates robust operations in noisy environments with low SNRs, which means that even with very low transmitted power (in order to reduce the risk of interfering with other users of the same band and also to keep the implant device small with long battery life), we can achieve a high localization accuracy. Among many
applications for in-body localization, we only make a brief description about two of them in the following subsections.

5.1.1. Wireless Capsule Endoscopy

Recently, Wireless Capsule Endoscopy (WCE) has become an admired method and preferred technique to diagnose and visualize the human gastrointestinal (GI) tract [49]. WCE has many advantages for physicians and patients compared to other in-body diagnostic methods because it is a non-invasive, highly accurate, and portable procedure [49], [50]. Unlike traditional endoscopic methods, such as colonoscopy and gastroscopy which can only reach the first or last several feet of the gastrointestinal tract, WCE is capable of visualizing the entire GI tract [50], [51]. Figure 5.1 shows two of the popular endoscopic capsule brands as well as an illustration for the size of the capsule [72], [73], [74], [75], [76].

During the process of capsule endoscopy, physicians need to know and recognize the precise position of the endoscopic capsule when each image is taken. Thus, accurate capsule localization becomes essential in capsule endoscopy [49], [50] and has been extensively investigated in the past few years [49]-[57]. Among the different methods that have been proposed for capsule positioning, radio frequency (RF) signal based techniques have the advantages of lower hardware costs and wider applicability (i.e., non-application specific) [49], [50], [51]. However, the RF based techniques need to be compliant with the Medical Implant Communication Services (MICS) regulations on signal bandwidth and power.
5.1.2. Tumor Localization and Tracking in Radiation Therapy

Radiation therapy (also called Radiotherapy) is an effective method to combat cancerous tumors by delivering high doses of radiation to the tumor to kill or control the growth of malignant cells by damaging its DNA [58]. A critical requirement of radiation therapy is to precisely delineate tumors and adjacent normal structures to avoid healthy tissues exposed under radiation. Recent techniques, such as three-dimensional conformal radiation therapy (3DCRT) and Intensity-Modulated Radiation Therapy (IMRT) [71], have significantly enhanced the ability to deliver an accurate radiation dose to the target volumes. In these methods, the radiation is split into hundreds of thin beamlets targeting the tumor from various angles to achieve a better focus on the cancerous region and reduce the damage to the surrounding healthy cells [59]. In IMRT, beamlets can also have various radiation intensities and it helps to produce a treatment area that better conforms to the contour of the tumor [60]. IMRT is an effective and efficient tool to treat the concave-shape tumors such as tumors wrapping around the organs.

Knowing the exact position of the tumor is one of the most essential prerequisites in radiation therapy, because any slight bias in the position of the tumor will cause the radiation to be delivered to the surrounding healthy tissues rather than the tumor area,
which not only degrades the performance of the treatment due to a lack of sufficient dose for the tumor treatment, but also it may cause severe side effects such as secondary cancer [59], [61]. It is important to note that the position of the tumor changes during radiation therapy because of respiration, gastro-intestinal, bladder filling, cardiac system or patient movements. Thus a real-time tumor tracking mechanism is highly desired in radiation therapy treatments in order to deliver and maintain a precise amount of radiation to the tumor region without damaging the surrounding healthy tissues [59], [61].

Various methods have been proposed in the literature for tumor tracking in radiation therapy treatments based on implanting several wired or wireless devices (called beacons) inside or in the vicinity of the tumor [59]-[70]. The Calypso Localization system is one of the most prevalent methods that has been widely used for tumor positioning in prostate radiation therapy [64], [65]. In the Calypso system, three magnetic transponders are implanted inside or in the vicinity of the target. Localization of the transponders is achieved using an electromagnetic array consisting of four electromagnetic coils to excite the transponders and 32 receiving coils to pick up the response from the transponders. Positions of the implanted transponders are estimated relative to the magnetic array based on the response measurements [64], [61]. Figure 5.2 demonstrates the Calypso system, console, electromagnetic detector array, and implants [61], [64], [65], [77]. There are several other electromagnetic tracking systems such as the methods proposed in [59], [70] that use the similar idea to track the tumor positions during the radiation therapy.

In this research, we propose and develop a novel positioning method to achieve very accurate results. The proposed system can be faster compared to common magnetic tracking methods, since there is no need for transponder excitation or for frequently
switching of the magnetic array between transmit and receive modes. In the proposed method, we use only one wireless implantable RF transmitter (that has the potential to be smaller than magnetic transponders because no RLC resonance circuit is needed) implanted inside or in the vicinity of the tumor.
Figure 5.2: The Calypso System, electromagnetic detector array and console, and implants [61], [64], [65], [77].
5.2. Challenges for In-Body Localization

Some serious challenges exist for in-body localization because of the complex structure of human body, which make the implant localization very difficult and very different compared to regular localization in other environments. In the following, we review four of the difficulties exist for in-body localization.

5.2.1. Signal Propagation Velocity

The human body is made up of various organs that consist of different types of tissues; thus, the electrical characteristics of the body such as conductivity and relative permittivity show significant heterogeneity. For example, the relative permittivity value varies according to the tissue type. Since the signal propagation velocity is expressed as a function of the relative permittivity, the propagation velocity and consequently the time-of-arrival (TOA) highly depends on the specific tissue layers that the signal passes through. In other words, in free space, we can reasonably assume that the signal propagation velocity is constant. However, this assumption will no longer be valid for the scenario inside the human body, and the propagation velocity varies through different body tissue types. Consequently, the delay (time-of-arrival) relies on the specific tissue that the signal passes through from the implant (emitter) to the sensor (receiver). This velocity inconsistency makes it very complicated to calculate or interpret the delay of the overall path from the implant to the sensors.
5.2.2. Signal Attenuation Parameters

The *path loss exponent* and *power absorption parameters* also vary by tissue thickness [78]. In other word, the path loss model for the signals propagating in narrow tissue (the distance less than 10 cm from body surface) is different from the propagation model for deep tissue (more than 10 cm). Therefore, traditional in-body localization methods based on RSS or TOA are inaccurate unless we have *a priori* knowledge about the position of the implant. Even if this a priori information is available it is difficult to exploit it in the classical location methods.

5.2.3. Severe Multipath Condition

The human body is made up of various organs and different types of tissues. One of the significant challenges in in-body signal propagation and wireless communication is the severe multipath problem caused by the signal reflections at boundaries of organs. The multipath issue can significantly degrade the quality and performance of classic localization methods (especially the TOA based methods). Thus, for in-body localization, the methods which are more robust to multipath conditions are highly desirable.

5.2.4. Power and Bandwidth Restrictions

The safety restrictions on signal power and bandwidth with regard to the protection of human health also make it more challenging to achieve accurate location estimation in implant in-body localizations. The Medical Implant Communication Services (MICS) standard regulates the two-way communications with medical implants to be in frequency band 401-406 MHz with the maximum transmitted power of 25 μw and the maximum
bandwidth of 300 KHz [50]. It is therefore desirable to pursue novel localization techniques that are effective and accurate under these constraints.
Chapter 6

Spatial Sparsity Based Medical Implant
In-Body Localization Using Wireless Body Sensor Network

6.1. Spatial Sparsity Based Approach for In-Body Localization

Wireless devices for medical diagnostics and therapeutics have gained increasing attention in the context of quality healthcare. Implantable and wearable medical devices, such as smart pills, body sensors, pacemakers, insulin pumps and etc. have been playing significant roles in healthcare systems by capturing, transmitting and controlling the vital information of patients.

In this research, we develop a novel method based on spatial sparsity for in-body localization using wireless body sensor network to achieve very precise results for medical implant positioning. In this method, the implant plays the role of an emitter by transmitting an RF signal. The signal will be received by a sensor array mounted in a known position (Figure 6.1). The sensors can be mounted on the body surface, inside a wearable jacket, or in any known position beneath or above the patient’s body. Then, the received signals will be processed to estimate the position of the implant (i.e., the
emitter), based on received signal strength (RSS) combined with time-of-arrival (TOA) and constant phase parameters.

Figure 6.1: (a) Implant and sensors configuration for three sensors. (b) Sensor array mounted on body surface or mounted inside a wearable jacket.

As mentioned in previous chapters, the classical localization methods usually include two stages; in the first stage, one or more location-dependant parameters (such as TOA or RSS) are estimated. Then in the second stage, these parameter estimates are used in statistical processing or finger printing methods to estimate the location of target. However the classic two-stage method is not necessarily optimal because in the first stage the parameter estimates are obtained by ignoring the fact that all measurements should be consistent with a single target location. In other words, each stage itself is itself optimal but the cascade of the two stages is not necessarily optimal [20], [24].
Unlike traditional RSS or TOA based localization approaches, we use convex optimization theory to solve the problem of location estimation directly without going through the intermediate stage of TOA or RSS estimation. In other words, there is no need to explicitly estimate the received signal strength or time-of-arrival for each one of the sensors by its own in a separate stage and the decision about the target location is made directly based on all received signals. Given all aforementioned features, the proposed method is much more accurate in positioning and very robust to multipath conditions caused by signal reflections at the boundaries of body organs compared to classic TOA/RSS based methods.

As mentioned in chapter 4, there are some significant and unique challenges for the localization procedure inside a human body compared to the localization problems in other environments. Two of the main challenges are the dependency of the signal propagation velocity on the specific tissue layer that the signal is passing through, and the dependency of the signal absorption parameters and model on the tissue thickness. However, in the proposed method, we exploit these characteristics as location-dependent parameters to achieve more accurate estimations.

We consider the human body area of interest as a fine enough three-dimensional grid space. In this case, the number of grid points containing the implant(s) (i.e., the emitters) is much smaller than the number of all grid points in the space. Assume that we allocate a positive number (such as one) to the grid points that include an emitter and assign zero to the rest of the grid points. Now, by arranging those numbers into a vector, we obtain a very sparse vector including only one (or more if we have more than one implant) non-zero elements. Since each element of this vector corresponds to a grid point in the grid
space, we are able to estimate the location of the implant by finding the position of the non-zero elements of the sparse vector.

In chapter 3, we saw that the $\ell_2$-norm minimization will not achieve the sparse solution because $\ell_2$-norm just measures the signal energy not the sparsity. Unfortunately, the $\ell_0$-norm minimization which obtains the sparse solution is an NP-hard non-convex optimization problem, thus we need to find an alternative solution. It is shown that $\ell_1$-norm minimization is a reasonable approximation for $\ell_0$-norm minimization problems; it achieves the sparse solution very well and it is a convex optimization problem that can be simplified to a linear program known as basis pursuit [31].

In our localization problem, we seek to find the sparse vector that maximizes the fit of the predicted received signals based on the estimated location and the true observed signals by sensors. In other words, we can estimate the location of wireless implant(s) by extracting the positions of those non-zero elements in the “sparsest” vector that satisfies the delay and path lost relationship between transmitted signals (by implant) and observed signals (by sensors). To this end, after formulating the problem in terms of the unknown sparse grid vector, we can estimate this vector by enforcing the sparsity using $\ell_1$-norm minimization on the grid vector, and at the same time minimizing the cost between expected received signals (that satisfy the delay and signal strength relationship) and actual received signals by forming a least square problem.

6.2. Signal model and parameters

Suppose that a wireless medical implant transmits a signal and $L$ sensors receive that signal. The complex baseband signal observed by the $l^{th}$ sensor is
\[ s_{r,l}(t) = \alpha_t e^{j\phi_t} s_l(t - \tau_t) + n_l(t) \] (6.1)

where \( s_l(t) \) is the transmitted signal, \( \tau_t \) is delay, \( \alpha_t \) is the real path attenuation, \( \phi_t = -2\pi f_c \tau_t \) is constant phase, \( f_c \) is the carrier frequency, and \( n_l(t) \) is the noise. The path loss model in dB is given by [78],

\[ P_{\text{Loss}}(d) = P_{\text{Loss}}(d_0) + 10\beta \log_{10}(d / d_0) + R, \quad d \geq d_0 \] (6.2)

where \( P_{\text{Loss}}(d) \) is the path loss at distance \( d \), \( P_{\text{Loss}}(d_0) \) is the path loss at the reference distance \( d_0 = 50 \text{ mm} \), \( \beta \) is the path loss exponent value and \( R \) is a zero mean Gaussian random variable (in dB) representing the shadowing effect, \( R \sim N(0,\sigma_R^2) \) [78]. Table 6.1 shows the path loss parameters for the path from implant to the body surface [78].

<table>
<thead>
<tr>
<th>Implant to Body Surface</th>
<th>( P_{\text{Loss}}(d_0) ) (dB)</th>
<th>( \beta )</th>
<th>( \sigma_R ) (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deep Tissue</td>
<td>47.14</td>
<td>4.26</td>
<td>7.85</td>
</tr>
<tr>
<td>Near Surface</td>
<td>49.81</td>
<td>4.22</td>
<td>6.81</td>
</tr>
</tbody>
</table>

We have to note that the term \( e^{j\phi_t} = e^{-j2\pi f_c \tau_t} \) includes important information about the signal phase and delay. This information can be very beneficial in location estimation, especially for narrowband signals where it is difficult or impossible to estimate the delay from the signal samples. Traditional TOA based localization methods usually don’t exploit the term \( \phi_t = -2\pi f_c \tau_t \) in location estimation, and that is why for narrowband
signals (such as the signals we use for in-body localization), the traditional methods fail to achieve accurate results. However, in the proposed method, since we use the entire signal to estimate the location directly, we take advantage of this term to achieve much more accurate results.

In free space, we can reasonably assume that the signal propagation velocity is constant. However, this assumption will no longer be eligible for the scenario inside the human body, and the propagation velocity varies through different body tissue types. Consequently, the delay (time-of-arrival) relies on the specific tissue that the signal passes through. The signal propagation velocity $v(\omega)$ at a specific frequency in a homogeneous tissue is given by,

$$ v = cVF(\omega) $$

$$ VF(\omega) = \frac{1}{\sqrt{\varepsilon_r(\omega)}} $$  \hspace{1cm} (6.3)

where $c$ is the speed of light in the free space, $VF(\omega)$ is velocity factor and $\varepsilon_r(\omega)$ is the relative permittivity of a human body tissue at frequency $\omega$. As we see in (6.3), the relative permittivity is frequency dependent. However, the curves and values of relative permittivity are available for various frequencies and different body tissues (such as fat, muscle, bone, lung, stomach, intestine, tendon and so on) [79], [80].

The overall signal delay on the path from the implant to the receiving sensor (after passing through $N_I$ different tissues) is calculated as following:
\[ \tau = \sum_{i=1}^{N_l} \tau_i = \sum_{i=1}^{N_l} \frac{d_i}{v_i} \]
\[ = \frac{1}{c} \sum_{i=1}^{N_l} d_i \sqrt{\varepsilon_i} \]
\[ = \frac{d}{c} \sum_{i=1}^{N_l} k_i \sqrt{\varepsilon_i} \quad (6.4) \]

where \( N_l \) is the number of different tissue layers on the path, \( \tau_i, d_i, v_i \) and \( \varepsilon_i \) are the delay, length, velocity and relative permittivity of \( i^{th} \) tissue respectively (at desired frequency) and \( k_i = (d_i / d) \) is the percentage of each tissue on the path. We can rephrase equation (6.4) as,

\[ \tau = \frac{d}{v_{eq}} \]
\[ v_{eq} = \frac{c}{\sum_{i=1}^{N_l} k_i \sqrt{\varepsilon_i}} \quad (6.5) \]

where \( v_{eq} \) is the equivalent propagation velocity. We seek to estimate the equivalent signal propagation velocity \( v_{eq} \) for the paths from each grid point to the sensors. This estimation is accomplished by calculating the percentage of each tissue layer on the entire path. Note that this velocity estimation can be performed in an off-line manner given the configuration of the body tissue layers, which could be acquired beforehand from MRI or CT systems [52]. Such propagation velocity information will be used later in the real-time implant localization and tracking step.

However, calculating the length of \( d_i \) and consequently the value of \( k_i \) in equation (6.5) is very challenging because the boundary surfaces of tissues are not mathematically determined. We need to compute \( d_i \) as the length of a portion of the line-of-sight that is limited between two tissue boundary surfaces. To this end, we are seeking an effective
algorithm to find the intersection points between each one of the tissue boundary surfaces and the line-of-sight.

We propose to find the intersection points between boundary surfaces and the line-of-sight (passing through from emitter to the sensor) by solving the following minimization problem over all grid points on each boundary surface:

\[ I = \arg \min_p (L_1 + L_2) \quad (6.6) \]

where \( I \) is the intersection point, \( L_1 \) is the length of the straight line from the sensor to the grid point \( p \) and \( L_2 \) is the length of the line from grid point \( p \) to the implant. As we see in Figure 6.2, by minimizing the length \( (L_1+L_2) \) the multisegment line SPE converges to the straight line-of-sight SIE and this minimization occurs at the intersection point \( I \).

![Figure 6.2: Minimization over all grid points \( p \) on the surface to find the minimum of \( (L_1+L_2) \) that happens at intersection point \( I \).](image)
After finding the intersection point that meets the minimization condition in (6.6) for each of the boundary surfaces, the length $d_i$ can be obtained for each tissue layer as the distance between the two intersection points. After obtaining all $d_i$’s for a specific path, we are able to calculate the $v_{eq}$ for that path using equation (6.5).

Figure 6.3 shows an example of tissue configuration consisting of three different tissue layers (fat, muscle and an organ tissue such as lung tissue) and their boundary surfaces among various tissue layers. In these figures, the top surface is the body skin, the middle surface is the boundary between muscle and fat and the bottom surface is the boundary between muscle and the organ tissue. The red points on the body skin are the sensors and the blue line is the line-of-sight from one of the sensors to the implant. $I_1$ and $I_2$ show the intersection points.
Figure 6.3: An example of tissue configuration from two different views. In both figures, the top surface is the body skin, the middle surface is the boundary between muscle and fat and the third surface is the boundary between muscle and the organ tissue. The red points on the body skin are the sensors and the blue line is the line-of-sight from one of the sensors to the emitter. $I_1$ and $I_2$ show the intersection points.
6.3. Problem Formulation

In this section, we present the problem formulation and show how we exploit the spatial sparsity to solve the location estimation problem. To clarify the formulations and distinguish the variables, we will use the lower-case italic bold letters to present a vector, upper-case italic bold letters to show a matrix and lower-case italic non-bold letters to present scalars.

Assume that each sensor collects $N_s$ signal samples at sampling frequency $F_s = 1/T_s$. Then we have

$$s_{r,l} = \alpha_l e^{j\phi_l} D_l s_t + n_l$$

(6.7)

where $(.)^T$ is the matrix or vector transpose, $s_{r,l}$ is the vector containing $N_s$ samples of the received signal by $l$th sensor, $s_t$ is the vector of $N_s$ samples of the transmitted signal, $n_l$ is the noise vector and $D_l$ is the time sample shift operator by $m_l = (\tau / T_s)$ samples where $\tau = (d_i / v_{eq,i})$ is delay, $d_i$ is the distance between implant and the $l$th sensor and $v_{eq,i}$ is the equivalent velocity on the path from emitter to the $l$th sensor derived from (6.5). We can write $D_l = D^m$ where $D$ is an $N_s \times N_s$ permutation matrix defined as,

$$[D]_{ij} = \begin{cases} 1 \text{ if } i = j + 1 \\ 1 \text{ if } i = 1, j = N_s \\ 0 \text{ otherwise} \end{cases} ; \quad D_l = \begin{bmatrix} 0 & 1 \\ \vdots & \vdots \\ 0 & 1 \\ 1 & 0 \end{bmatrix}^m$$

(6.8)
Now, we assign a number $\delta_{x,y,z}$ to each one of the grid points $(x,y,z)$. Assume that $\delta_{x,y,z} = 1$ for the grid point containing the implant beacon and $\delta_{x,y,z} = 0$ for the rest of the grid points. Thus, by rephrasing the equation (6.7), the signal vector received by $l^{th}$ sensor is given by,

$$
\mathbf{s}_{r,l} = \sum_{x} \sum_{y} \sum_{z} \delta_{x,y,z} \alpha_{l,x,y,z} e^{j\phi_{l,x,y,z}} d_{l,x,y,z} \mathbf{s}_t + \mathbf{n}_l
$$

(6.9)

where $d_{l,x,y,z}$ is the time sample shift operator with respect to sensor $l$ assuming that the emitter is located in the grid point $(x,y,z)$ and the summations are over all grid points in the desired $(x,y,z)$ range. Note that $d_{l,x,y,z}$ and $\alpha_{l,x,y,z}$ are known in (6.9) because the position of the sensor $l$ and each grid point $(x,y,z)$ is known and thus we are able to find the path loss and delay from equations (6.2) and (6.5) for the path from grid point $(x,y,z)$ to the sensor $l$. The unknown term is $\delta_{x,y,z}$ which represent whether a grid point contains the implant or not. If we reform all grid points into a column vector and re-arrange the indices, then we will have

$$
\mathbf{s}_{r,l} = \sum_{n=1}^{N} \delta_{n} \alpha_{l,n} e^{j\phi_{l,n}} d_{l,n} \mathbf{s}_t + \mathbf{n}_l.
$$

(6.10)

where $N$ is the total number of grid points. Now, we define the tall matrix $T_n$ as the delay and path-loss operator with respect to all $L$ sensors, assuming that the implant is located in grid point $n$ (the received signal comes from the grid point $n$) as following:
\[
\mathbf{T}_n \triangleq \left[ \begin{array}{c}
\alpha_{1,n} e^{i\phi_{1,n}} D_{1,n} \\
\alpha_{2,n} e^{i\phi_{2,n}} D_{2,n} \\
\vdots \\
\alpha_{L,n} e^{i\phi_{L,n}} D_{L,n}
\end{array} \right]_{L N_s \times N_s}
\]

where \( N_s \) is the number of signal samples.

Then, we define \( \boldsymbol{\theta}_n, n \in \{1, 2, \ldots, N\} \) as an \( LN_s \times 1 \) vector including all signals received by all \( L \) sensors when the emitter is located at grid point \( n \) as,

\[
\boldsymbol{\theta}_n = \mathbf{T}_n \times \boldsymbol{s}_r
\]

where \( \times \) is the regular matrix multiplication.

Now, we expand our results over all grid points by arranging all vectors \( \boldsymbol{\theta}_n \) for \( n:1 \ldots N \) as the columns of a matrix \( \mathbf{\Theta} \) as,

\[
\mathbf{\Theta} = [\boldsymbol{\theta}_1 \quad \boldsymbol{\theta}_2 \quad \ldots \quad \boldsymbol{\theta}_N]_{LN_s \times N}
\]

then we have

\[
\boldsymbol{s}_r = \mathbf{\Theta} \times \boldsymbol{v} + \boldsymbol{n}
\]

where \( \boldsymbol{s}_r = [\boldsymbol{s}_{r,1}^T \quad \boldsymbol{s}_{r,2}^T \quad \ldots \quad \boldsymbol{s}_{r,L}^T]^T \) is the \( LN_s \times 1 \) vector of all received signal samples by \( L \) sensors, \( \boldsymbol{v} = [\delta_1 \quad \delta_2 \quad \ldots \quad \delta_N]^T \) is the sparse vector of \( \delta \)-values assigned to each grid point and \( \boldsymbol{n} \) is the vector of noise. Since each element of vector \( \boldsymbol{v} \) corresponds to a grid point in the grid space, we can estimate the location of the implant by extracting the position of non-zero elements of the sparsest vector \( \boldsymbol{v} \) that minimizes the cost (fit of
solution to the observation) which is defined as the Euclidean distance between \( s_r \) and \( \Theta \times v \).

Thus, we can solve our problem by forming the regularized Basis Pursuit Denoising (BPD or equivalently Lasso) problem as [32]:

\[
\hat{v} = \arg \min_v (\|v\|_1 + \lambda \|\Theta \times v - r\|_2)
\]

(6.15)

where \( \| \cdot \|_p \) is the \( \ell_p \)-norm defined as \( \| v \|_p = \left( \sum |v_i|^p \right)^{1/p} \) and \( \lambda \) is the regularization parameter that adjust the balance between cost (fit) and sparsity. Selecting a proper \( \lambda \) is important in regularization problems; by increasing the value of \( \lambda \), the penalty on cost gets higher and it leads to a good fit to the observation. However, if we select a too large regularization parameter, it may fail to achieve a sparse solution. Various methods have been proposed in the literature for automated \( \lambda \) selection (such as in [33]). However, it is usually possible to fine reasonable choices for \( \lambda \) by empirical methods.

### 6.4. Reducing Computational Complexity

In this section we propose an alternative approach to reduce the computational complexity of the problem. This is very helpful in cases when we have a big number of grid points, such as implementing 3-dimensional high-resolution localization in a large area. In this method, we can split the area of interest into several sub-areas. We can form (6.15) for each sub-area and find the best grid point that meets the minimization problem. In a second step, we will only form the minimization problem among the selected grid points from each sub-area to find the final result. Note that the computational complexity of the minimization problem in (6.15) using traditional convex optimization methods for
a vector $\mathbf{v}$ of length $N$, is about $O(N^3)$. Thus, the computational complexity decreases significantly by reducing the size of the sparse vector.

Assume that we aim to estimate the implant position in a $N_x \times N_y \times N_z$ grid space. Thus, the vector $\mathbf{v}$ has $N_x N_y N_z$ elements and the computational complexity of the minimization problem is $O((N_x N_y N_z)^3)$. Now, if we split the grid space into $N_z$ grid planes with size of $N_x \times N_y$ each (as shown in Figure 6.4), we will have the complexity of $O((N_x N_y)^3)$ for each plane. Consequently, the total computational complexity will be $O(N_z (N_x N_y)^3)$ for all $N_z$ grid planes in addition to $O(N_z^3)$ for final step. This is almost $O(N_z^2)$ times smaller than the original problem.

![Figure 6.4: Splitting the 3D grid space into 2D grid planes](image)

Another method that can be applied to reduce the computational load is to start with a coarse grid, and then increase the resolution of the grid iteratively. In other words, after estimating the approximate location of the target in a coarse grid space, we generate a finer grid in a smaller area of interest around the approximately estimated position, and then re-calculate the location in this new area. We continue this process until we achieve the desired positioning accuracy.
Furthermore, to reduce the computational complexity and achieve more accurate results in real-time implant tracking, we can exploit the prior trajectory information, to estimate the new position of the implant. In this case, we only need to start the localization process over the entire grid space once to estimate the initial position. After that, we can always restrict the area of interest to only a small vicinity around the past estimated position, considering the maximum possible movement of the subject. This technique leads to a significant reduction in size of the sparse vector and computational complexity, and also helps to achieve more accurate results by choosing a finer grid in the specific area of interest.

6.5. Simulation Results

We evaluated the performance and accuracy of the proposed method using various sets of Monte-Carlo computer simulations for both endoscopic capsule localization and tumor tracking. In tumor tracking simulations, we used more number of signal samples for location estimation as well as finer grid (10 times finer than the grid used for endoscopic capsule) because we aimed to achieve more accurate results for tumor tracking case (of course with the cost of higher computational complexity).

The results demonstrate that the proposed method is very accurate in location estimation even when we use a small number of sensors (only 4 sensors) and a small number of signal samples (only 64 signal samples). Furthermore, the system shows robust operations and high performance in noisy environments (low SNRs). It means that we are able to achieve high localization accuracy even with very low wireless transmitted
power which helps to reduce the size of the implant device, increase the implant’s battery life, and also reduce the risk of interfering with other users of the same band.

In subsection 6.5.2, we also consider the case when the body tissue configuration is not perfectly known and there is uncertainty on the tissue configuration and tissue boundaries. We will see that the proposed method can achieve very accurate results even in the case when there is uncertainty about the tissue configuration (a boundary variation level of 3mm only increased the localization error by 0.5mm).

### 6.5.1. Simulation Results for Capsule Endoscopy

In the first scenario, we ran the Monte Carlo simulation for various numbers of receiving sensors (4, 8, 12 and 16 sensors) and various SNR values (0dB, 10dB and 20dB) in multipath conditions (100 runs each time). The random variables $R$ in (6.2) are generated with the parameters available in Table 6.1. In this simulation, the sensors are mounted on the human body surface with the configuration showed in Figure 6.5. The locations of the implant and the reflector points have been chosen randomly (the implant and reflector locations are not necessarily on grid points. The term reflectors refer to the points that reflect the transmitted signal and lead to multipath problem. The signal reflection usually happens at boundaries of organs). In this simulation, the signals are BPSK modulated, with frequency of 406 MHz and bandwidth of 300 KHz. Each sensor collects only 64 signal samples and the grid size is $(20 \times 20) \text{cm}$ with grid point spacing of 1 cm.
Figure 6.5: Sensors location and configuration on the body skin for 4, 8, 12 and 16 sensors. The scale units are in cm.

Figure 6.6 shows the RMS Error versus number of sensors for capsule localization in the (x,y) plane according to the sensor topologies shown in figure 6.5. Consistent with our expectation, the accuracy is dramatically improved by increasing the number of the sensors. Meanwhile, the results demonstrate that the proposed method can achieve a very high performance even for the low SNR cases. As we will see in tumor tracking simulations, it is even possible to achieve higher accuracy by increasing the number of observed signal samples and using finer grids (reducing the grid cell size).
In the second scenario, we used 16 sensors to estimate the 3D location of a moving endoscopic capsule in a 3-dimensional body space (Monte Carlo simulation with 100 runs each time). We used a noisy BPSK signal with frequency of 406 MHz and bandwidth of 300 KHz and SNR=20dB. In this simulation, we also simulated the
multipath condition (signal reflections from tissue boundaries) using reflector points at randomly chosen locations.

Figure 6.7-(a) demonstrates the true position of the implant (in blue) and the estimated position by the proposed method (in red). Figure 6.7-(b) shows the overall RMS error in positioning defined as

$$ e = \sqrt{e_x^2 + e_y^2 + e_z^2} $$

(6.16)

where $e_x$, $e_y$, and $e_z$ are RMS errors in positioning for X, Y and Z dimensions. It is shown that the positioning error is less than 7 mm, in the worst case. Note that the position of the implant is not necessarily on the grid points. In this case, the algorithm finds the nearest grid point to the implant location. Thus, a significant portion of the errors are resulted from the distance between the implant and the determined nearest grid point close to it. This kind of error can be reduced by employing finer grids in the processing, which would of course increase the computational overhead accordingly.
Figure 6.7: (a) True position of the endoscopy capsule (in blue) and the estimated position. (b) Error in positioning for each location in part (a).
6.5.2. Simulation Results for Tumor Localization and Tracking

In tumor tracking simulation, we examined the performance of the proposed method in the case when the body tissue configuration is not exactly known and there is uncertainty on the tissue configuration and tissue boundaries. Figure 6.8 shows an example when the signal (emitted by the target implant) travels through three different tissues including organ tissue (such as lung tissue), muscle and fat to reach the external sensors. In this case the sensors are mounted on the body surface, although they can also be mounted anywhere above or beneath the patient’s body. Figure 6.8-(a) shows the case when the tissue configuration is exactly known and Figure 6.8-(b) illustrates the case when a certain level of uncertainty of tissue boundaries exists. To emulate this uncertainty, we added unknown random noise to the simulated, exactly determined tissue boundaries. In this simulation, we used 16 sensors to receive a noisy BPSK signal (zero-mean white Gaussian noise) with SNR=20dB. Each sensor collects 256 signal samples, and the grid size is $5\times5$ cm with grid point spacing of 1 mm.
Figure 6.8: A sample tissue configuration. The top surface is the body skin, the middle surface is the boundary between fat and muscle and the bottom surface is the organ surface (such as lung surface) which is the boundary between muscle and organ tissue. The red points on the body skin are the sensors and the blue line is the line-of-sight from one of the sensors to the emitter. (a) perfectly known boundaries (b) approximately known boundaries generated by adding unknown noise to the boundary surfaces.
Figure 6.9-(a) shows a simple pattern for tumor (or implant) movement in the X direction. Figure 6.9-(b) shows the RMS (Root Mean Square) Errors of the implant localization results for the aforementioned two scenarios (Monte Carlo simulation results with 100 runs each time). As we see, the simulation results show the accurate localization and high performance of this method. The localization error is less than 1.5mm when the tissue configuration is exactly known and less than 2mm for underdetermined tissue boundaries. It is worth highlighting that a boundary variation level of 3mm only increased the localization error by 0.5mm. Note that the position of the implant is not necessarily on the grid points. In this case, the algorithm finds the nearest grid point to the implant location.
Figure 6.9: (a) A simple pattern for tumor (or implant) movement in X direction, (b) RMS Error for implant location estimation for the movement pattern in (a) in two cases: when the tissue boundary surfaces are perfectly known (red curve) and when there is some uncertainty about the boundary surfaces (blue curve).
Chapter 7
Conclusion and Future Work

In this dissertation, we developed a novel and effective method to estimate the position of medical implants using wireless sensor network. In this method we use both received signal strength (RSS) and time-of-arrival (TOA) for location estimation based on sparsity of the implant in three-dimensional space. Some unique challenges exist for in-body localization of a medical implant due to the complex nature within the human body, such as the dependency of the propagation velocity on type of the tissues, the dependency of the parameters of the path loss model on tissue thickness (deep or near surface), and the multipath problem caused by signal reflections at organ boundaries. Furthermore, the safety restrictions on signal power and bandwidth with regard to the protection of human health also make it more difficult to achieve accurate implant location estimation.

In this research, we developed a novel tissue-adaptive method by considering both signal propagation velocity and path loss exponent as location-dependent parameters that can be exploited to estimate the implant location more precisely. Unlike the classic methods, we directly estimate the implant location without going through the
intermediate stage of RSS or TOA estimations. In this method, we estimate the equivalent signal propagation velocity as well as the path loss model parameters for the path from each one of the grid points to the receiving sensors. After formulating the localization model in terms of the unknown sparse grid vector, we estimate that vector by enforcing the sparsity, and at the same time minimizing the cost between predicted received signals and actual received signals by forming a BPIC or BPDN problem.

The simulation results demonstrate that the proposed method is very accurate in location estimation even when we use a small number of sensors (only 4 sensors) and a small number of signal samples (only 64 signal samples). Furthermore, the system shows robust operations and high performance in noisy environments (low SNRs). It means that we are able to achieve high localization accuracy even with very low wireless transmitted power which helps to reduce the size of the implant device, increase the implant’s battery life, and also reduce the risk of interfering with other users of the same band.

We also applied the proposed sparsity based method for other localization scenarios including indoor patient localization for assistive healthcare and fall detection, and also emitter localization using fast moving sensors. Since the emitter location is estimated based on the original received signals altogether, rather than the parameters derived individually from each signal, the proposed method shows better performance compared to two-stage methods. Furthermore, the effect of multipath will be degraded in this method, because the indirect signal reflections are usually not consistent with a single location. Thus, the multipath signals combine incoherently in signal estimation. Pushing the sparsity in position domain also helps to reduce multipath effect, and make the algorithm capable to distinguish between locations of various emitters in multiple-emitter
scenarios. We compared the performance of the proposed method to two-stage classic localization method, as well as another one-stage method named direct position determination (DPD). We ran Monte Carlo simulations for different scenarios including single-emitter and multiple-emitters cases with or without multipath reflections. The Simulation results show the high performance and accuracy of the proposed method (especially in multipath conditions and multiple-emitters cases) compared to other methods.

Besides the implementations and evaluations that have been performed in this dissertation for implant in-body localization, there are still some more details to be considered in the future to achieve more realistic and accurate results. Although the computer simulations demonstrate the high performance and accuracy of the proposed method for capsule localization and tumor tracking, these methods should be evaluated in a more realistic platform (i.e. human body using real implant and data) before entering to practical and commercial stage.

This dissertation also inspires a number of new research directions. For example, as a future work, we can explore applying the proposed in-body localization method in EEG source localization, which is a long-standing and challenging problem. Similarly, we can explore novel application-based sparse signal processing algorithms, compressive sensing methods and data compression techniques for biomedical signals in different levels of body sensor networks (either sensor level or network level) for the purpose of transmission and computation load reduction.
References


