Localization with Multi-Lateration Techniques in Low-Power Wide-Area Networks

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Electrical and Computer Engineering
University of Saskatchewan
Saskatoon

By

Nhat H. Pham

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Disclaimer

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Electrical and Computer Engineering
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5A9 Canada

OR

Dean
College of Graduate and Postdoctoral Studies
University of Saskatchewan
116 Thorvaldson Building, 110 Science Place
Saskatoon, Saskatchewan, S7N 5C9 Canada

Abstract

This thesis concerns the localization of an Internet-of-Things (IoT) target (sensor) that intermittently transmits a signal (message) to an array of spatially separated receivers (gateways). In the setting of primary interest the source's signal is known to the receivers. The application of primary interest is a target (sensor) transmits a short time duration (burst) signal in a homogeneous line-of-sight environment to gateways that can measure the time of arrivals (ToAs) of the source's signal subject to suffering zero-mean Gaussian measurement errors. The time when the source emits the signal (i.e., the transmit time) is unknown. However the results can be applied in a wide range of applications.

The problem of localizing a target is approached from a probability density point of view as opposed to the well-explored point estimation point of view. Equations for the joint *a posteriori* probability density functions (pdfs) of the target's coordinates in both two-dimensional (2D) and three-dimensional (3D) spaces are developed using the measurements of the times that a transmitted message arrives at the spatially separated gateways. The *a posteriori* pdfs also take uncertainty, i.e., measurement errors, in the positions of the gateways into account.

The corroboration and utility of the *a posteriori* pdf are explored through various examples to demonstrate its superiority and usefulness. First, it is shown that the joint *a posteriori* pdf of the target's location is not always approximately jointly Gaussian, especially when the target is in close proximity to one of the gateways, or when the gateways are located in close proximity to each other. Next, several examples are provided to examine the effects of the measurement errors in the positions of the gateways to the spread and position of the resulting *a posteriori* pdfs. These examples also incorporates the *a priori* pdf, which makes the *a posteriori* pdf very useful in many practical scenarios, especially when the *a priori* pdf is quite restrictive. Lastly, an improvement of the *a posteriori* pdf is thoroughly analyzed for a scenario where there are multiple transmissions of the message from a target located at a fixed position.

While intermittently active targets are the rule in the application of primary interest,

a chapter is dedicated to targets that transmit a continuous signal that is unknown to the receivers. In a scenario where the transmitted signal is unknown to the receivers, the ToAs can not be estimated, but the time-difference-of-arrivals (TDoAs) can be estimated by time-windowing the continuous signal and cross-correlating the windowed segments. An *a posteriori* pdf is developed for TDoAs obtained by taking the difference of ToAs. It is shown that this *a posteriori* pdf does not apply to TDoAs obtained by cross-correlation unless one of the segments being cross-correlated is noise free.

Acknowledgments

It is a pleasure to thank many people during my studies who made this dissertation possible. A few words mentioned here cannot adequately express my appreciation.

First, I would like to express my deepest gratitude to my supervisors, Professor Ha H. Nguyen and Professor Eric J. Salt, whose expertise, understanding and patience throughout my research program added considerably to my graduate experience. Without them, this thesis would not have been completed.

Second, I would like to also thank the other members of the committee, Professors Rajesh Karki, Khan A Wahid, Ebrahim Bedeer Mohamed, Travis Wiens from the University of Saskatchewan, and Professor Henry Leung from the University of Calgary for reviewing and evaluating this thesis. Their insightful comments and suggestions have significantly improved the quality of this thesis.

Last but not least, I would like to thank my family for the support that they have provided me throughout my studies. Without their love and encouragement, I would not have finished this thesis. My special thanks are extended to all my friends: Binh, Long, Shania, Ali, Botao, Khai and Nghia in Communications Systems Research Group (CSRG) for sharing their knowledge and invaluable assistance.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2D	2-Dimensional
3D	3-Dimensional
C/A	Coarse Acquisition
UTC	Coordinated Universal Time
GDoP	Geometric Dilution of Precision
GPS	Global Positioning System
HDOP	Horizontal Dilution of Precision
IoT	Internet of Things
LoS	Line of Sight
LLS	Linear Least Squares
LoRaWAN	Long Range Wide Area Network
LPWANs	Low-Power Wide Area wireless Networks
NLS	Non-linear Least Squares
NLoS	None Line of Sight
MAP	Maximum *A Posteriori*
ML	Maximum Likelihood
MMSE	Minimum Mean Squared Error
PDoP	Position Dilution of Precision
PDF	Probability Density Function
PRN	Pseudo-Random Noise
RF	Radio Frequency
RV	Random Variable
RSS	Received Signal Strength
SOCP	Second-Order Cone Programing
SDP	Semi-Definite Programming
SNR	Signal to Noise Ratio
ToA	Time of Arrival
ToF	Time of Flight
TDoA	Time Difference of Arrival

TCP	Transmission Control Protocol
UDP	User Datagram Protocol
WLLS	Weighted Linear Least Square
VDOP	Vertical Dilution of Precision

1. Introduction

There is a variety of applications where one needs to know the location of an object. For example many applications that run on a cellular telephone need to know the location of the cell phone. Another example is trucking companies need to know the on-route locations of their trucks so that the people who load a truck can be scheduled to be on site when the truck arrives. Another example is locating house pets that run away from a transmitter located in the pet's collar. Also, our military depends heavily on locating objects. For example they would like to know the location of all the ships and submarines in the ocean all the time.

This thesis is concerned with the localization of low-power asynchronous burst transmissions. The transmitters would be placed on trucks or in the collars of pets or on/in one of many other things. Of interest in this thesis is *a posteriori* probability density function for the co-ordinates of the transmitter. The algorithm used to generate the *a posteriori* use the times the burst signal arrives at a network of spatially separated receivers. Of primary interest is a low-power wide-area network known as LoRaWAN and the transmitters designed to work within that standard.

1.1 A General LoRaWAN System

Low-power wide area wireless networks (LPWANs) are composed of battery-powered devices that communicate over long ranges at low bit rates with several centralized receivers that are connected to the Internet. Such networks allow data generated by remote sensors to be collected by an application server and used to control machines or initiate a human action. For example, temperature sensors that transmit a message when the temperature drops below 1°C could be deployed throughout an orange orchard. Upon receiving "temperature"

messages, the application server could send messages to turn on sprinklers that shower the orange trees providing extra thermal mass that prevents the trees from freezing.

Low-power wide area wireless networks are self-contained and do not have to adhere to standards. However, standards have been developed by service providers and manufacturers to allow interoperability of equipment from different manufacturers and, at the same time, increase the scale of manufacturing.

Example standards developed for LPWANs include Sigfox [1], Ingenu [2], DASH7 [3], LoRaWAN [4], and NB-IoT [5]. While different standards use different technologies and have their own advantages and disadvantages [6], this PhD research is concerned with LoRaWAN. This technology is gaining tremendous commercial growth in more than 100 countries around the world (see `https://lora-alliance.org/`). Furthermore, the industrial partner of this research, Cisco Systems Canada, is also a key member in the LoRa Alliance, an organization responsible in developing the LoRaWAN standard and promoting it.

A generic LoRaWAN network is shown in Figure 1.1. The network consists of 4 different types of equipment: (i) battery powered *sensors*, (ii) the sensor-internet interface referred to as the *gateways*, (iii) a computer that controls/administers the gateways referred to as the *network server*, (iv) a computer that interprets and reacts to sensor data referred to as the *application server* and (v) a computer to communicate with the application server referred to as the *e-commerce server*.

The purpose of the system is to gather information collected by a network of sensors and use this information to control them or possibly some other sensors. The information collected by the sensor is organized into a message and transmitted wirelessly over the ISM band (902 to 928 MHz in North America) to all gateways that are in range. All gateways that receive the message re-format it and send it over the Internet using User Datagram Protocol (UDP) to a single network server. The UDP protocol is a low-reliability "best effort" protocol that does not guarantee delivery of the message.

The network server analyzes the messages it receives to get the identification number of the sensor that originated the messages. The network server then forwards the messages

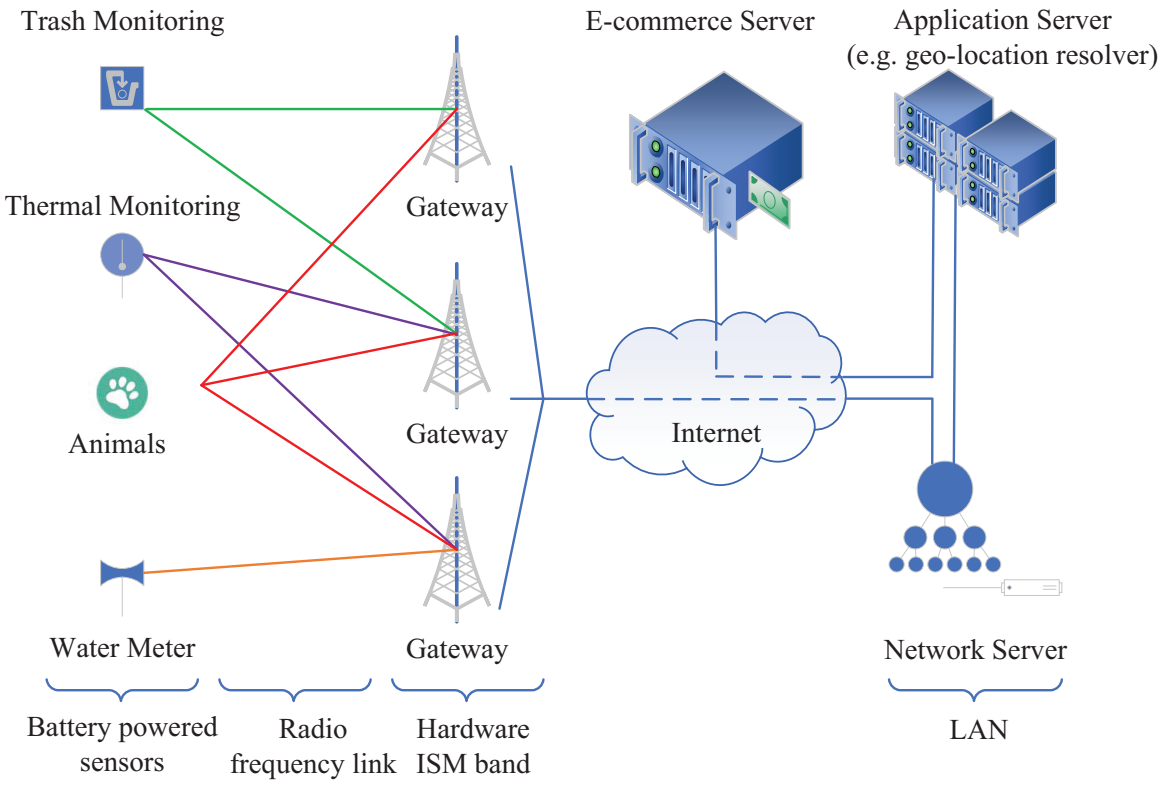

Figure 1.1: A generic LoRa network.

to the appropriate application server using the more reliable transmission control protocol (TCP).

The application server processes the messages, perhaps displaying information on a web-page or perhaps sending a control command through the network server and through a specified gateway to a network of sensors. In addition, the application server receives and processes requests from e-commerce servers initiated by end-users to serve their specific needs.

Commands that control sensors are originated from the application server. They are sent to the network server where they are forwarded to the gateways. The gateways broadcast the messages over the ISM band to the network of sensors.

The LoRaWAN standard specifies the radio frequency (RF) signals that link the gateways and sensors. The standard also specifies the interface protocols for communication between the application server and network server, network server and gateways, and gateways and

sensors. The only protocol of importance to this proposal is the protocol for communication between gateways and sensors. The standard allows for three different protocols that support different levels of trade-off between battery efficiency and gateway-to-sensor communication.

The class A protocol was derived to preserve battery in the sensors. In this protocol, the sensors transmit whenever they need to, which is usually not often, and have their receivers turned off except for a small window of time immediately after they transmit. While the time of transmission is determined by the sensor, the standard limits the transmission rate to something like 3 messages an hour. Limiting communication with a sensor to a short time immediately after it has initiated a message severely limits the ability of the application server to control the sensors. This is the price that must be paid to preserve battery power that would be consumed by an active receiver.

The class B protocol has the sensors communicating with the gateways at regularly scheduled times. This gives the application server better control over the sensors, but at the expense of battery power needed to run the receiver in the sensor.

The class C protocol is derived to give the application server very good control over the the sensors. In this protocol the sensors have their receivers turned on at all times. Of course this will be a steady drain on the battery. Practically, class C sensors are powered from the electrical grid.

The objective of this research is to find algorithms that run in the application server that not only estimate the location of Class A sensors but also determine the probability that the target is or is not in a specified region. There are two types of inputs. One is the geo-location of the gateways. The other is the "time stamp" inserted in the message by the gateways. The precision of both affect the precision of the estimate and therefore is discussed in the next section.

1.2 Timestamp Precision of a LoRaWAN Gateway

This section has been included for two reasons:

1. It is prudent to explain the physics that provides the geo-location of the gateways and the *time* used in the time stamp since the ability to locate a class A device depends entirely on the ac-curacies of the geo-locations of the gateways and the time stamp inserted by the gateways.

2. The methodology used to locate and set the time in a GPS receiver have been adapted to the localization of class A devices and provides background for several methods proposed in the literature. Some of these methods will be discussed later in this thesis and the background provided in this section provides a foundation.

It is pointed out that, while quite interesting and supportive of existing methods, this section, i.e., Section 1.2 does not directly contribute to the understanding of the techniques proposed n this thesis, and does not need to be fully understood or even read for the that matter, to understand the research reformed in this thesis.

The LoRaWAN standard supports the gateways "time stamping" messages with the times the messages are received before forwarding them to the network server. This is done by reformatting the message to include a field that contains the time the message was received from the sensor by the gateways. The gateways derive the time that they stamp in the messages from GPS satellites.

The time stamps in the messages as well as the geo-locations of the gateways are used in the application server to estimate the geo-location of the sensor that transmitted the messages. The error in the estimate of the geo-location of the sensor depends on the errors in times that are stamped in the messages as well as the errors in the geo-locations of the gateways. The gateways derive their geo-locations and the time used in the time stamps from the GPS satellites. To further understand the precision of the "time stamps" in the gateways, a basic understanding of the Global Positioning System is required.

A constellation of 27 orbiting satellites is used by the Global Positioning System (GPS) in order to provide positioning, navigation, and timing service [7]. They are arranged into six equally-spaced planes surrounding the Earth at the height of 20,200 km. At least 4 satellites are in the view of any point on the planet. The on-board oscillators of the GPS satellites,

which are made of either rubidium or cesium, are automatically adjusted to Coordinated Universal Time (UTC) managed by the United States Naval Observatory. Therefore, the GPS satellites become the primary source for time and frequency to most points on the planet. This time is distributed by broadcasting messages with an embedded pseudo-random noise (PRN) code, which is either a coarse acquisition (C/A) code with a chip rate of 1023 chips per millisecond or a precision (P) code with a chip rate of 10230 chips per millisecond. The PRN code is used to identify the broadcasting satellite [8]. The C/A code embedded in the message is transmitted in L-band on a carrier of frequency 1575.42 MHz (L1) and the P code embedded in the message is transmitted in L-band on two carriers of frequencies 1575.42 MHz and 1227.6 MHz (L2).

A gateway locates itself using messages broadcasted from GPS satellites. To simplify the explanation of how this is done, assume that the LoRaWAN gateways and the satellites are in the same plane. The location of Satellite i, $(x_i^{(s)}, y_i^{(s)})$ and the time at which the message is transmitted, $t_i^{(s)}$, are obtained in the message sent from Satellite i. The time the message from Satellite i arrives at the gateway according to the onboard-clock is denoted $\widehat{t_i}$. This time includes a bias b from the gateway's clock since its clock is not synchronized to the clock in the satellite. Since all the satellites are synchronized in time, the bias b is the same for all satellites. The Time-of-Flight (ToF) from Satellite i to the gateway is $(\widehat{t_i} - b - t_i^{(s)})$. For L satellites, the position (x_g, y_g) of the gateway is the solution to the set of range equations:

$$r_i = (\widehat{t_i} - b - t_i^{(s)})c \quad , \quad i = 1, 2, \ldots, L, \tag{1.1}$$

where $c = 2.9979 \times 10^8$ m/s is the speed of light and $r_i = \sqrt{(x_g - x_i^{(s)})^2 + (y_g - y_i^{(s)})^2}$ is the range from the gateway to Satellite i.

Several methods can be used to solve the set of equations generated in (1.1) and can be visually explained by their geometric interpretation. Let $\widehat{r_i} = r_i + b \times c$, $i = 1, 2, \ldots, L$, denote the measured range from the gateway to Satellite i. The most popular method estimates the position of the gateway as one point in the polygonal whose vertices are the intersections of pair-wise circles which are centered at $(x_i^{(s)}, y_i^{(s)})$ with radii $\widehat{r_i}$ as shown in Figure 1.2.

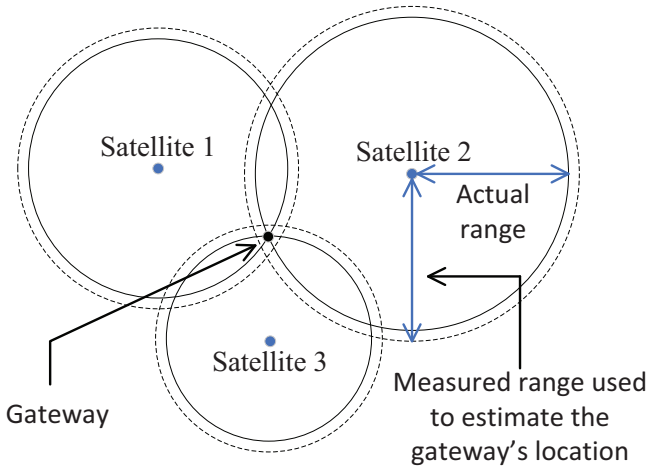

Figure 1.2: Circle geometry.

The exact position of the gateway coincides with the intersection of the circles with the correct ranges.

Another popular method uses one measurement to cancels out the bias b by using the measured range difference, denoted $\widehat{d}_{i,j} = \widehat{r}_i - \widehat{r}_j = r_i - r_j$. The locus of the points which have a constant difference between ranges from them to two satellites is a hyperbola. The position of the gateway is the intersection of two hyperbolas associated with two measured range differences. While it is possible to determine the location of the gateway within an order of centimeters, doing so is expensive. Commercially available LoRaWAN gateways [9] specify self-location accuracy of a few meters.

The GPS satellites can be also used to synchronize gateways in the network to minimize the relative errors in the timestamps. There are three techniques used to synchronize the clocks among gateways: one way technique, the common-view technique, and the carrier phase technique [8]. In the one-way technique, the gateway's clock is synchronized with the GPS satellite's clock by using the 1 pulse per second signal generated by the GPS satellite. The pulse controls a time interval counter, that estimates the propagation delay based on the co-ordinates of the satellite and gateway and the time sent by the satellite. A gateway that uses the C/A code can synchronize its time with an error that has a standard deviation of 10 ns or less [8].

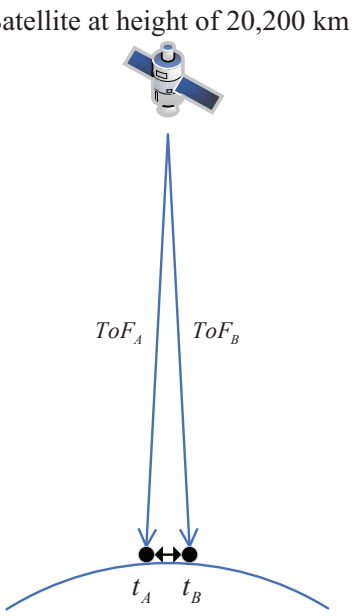

Figure 1.3: Common-view time transfer technique.

The common-view technique is an enhancement of the one-way technique designed to synchronize the clocks in two gateways. It is essential for the two gateways to receive the message sent from a common satellite at the same time. In order to meet this condition, the two gateways are scheduled to simultaneously observe the message sent from the common satellite based on the tracking schedules of the GPS satellites published by the Bureau International des Poids et Mesures. This technique works well providing the times the message arrives at the two gateways are nearly the same. Figure 1.3 shows an example of synchronizing the clocks of Gateways A and B in which the Time-of-Flight (ToF) of the messages propagating from the common satellite to Gateways A and B are ToF_A and ToF_B, respectively. The times the signal transmitted from the satellite arrives at Gateways A and B are recorded using the time of the local clocks. These times are denoted t_A and t_B, respectively. The difference between the clocks of Gateways A and B is $t_A - t_B - (\mathsf{ToF}_A - \mathsf{ToF}_B)$. The quantities t_A, t_B, ToF_A, and ToF_B are exchanged between the two gateways in order to set the clock of one gateway to be the same as the other. While most of the systematic errors (common errors) in the ToFs and arrival times at the two gateways completely cancel out, the independent errors, which are caused by things like the Johnson noise or local oscillators in the two gateways do not cancel out and impact the

accuracy of the time synchronization. The difference in time between two gateways after they are supposedly synchronized can have a standard deviation as little as 2.5 ns if the synchronization procedure utilizes measurements taken over 1-day. A standard deviation of 2.5 ns requires multiple satellites. With just one reference satellite, a standard deviation of 5 ns is possible [8].

The carrier phase technique is the enhancement of either the real-time one-way technique or common-view technique. It uses both the L1 and L2 carrier frequencies, which are 1000 times higher than the C/A code frequency in order to achieve higher resolution. The error in time after synchronizing using measurements taken over 1-day has a standard deviation of 250 ps or less. The drawback of this technique is that it requires extensive post-processing of the collected data and, hence, is more suitable for establishing exact time in a laboratory setting [8].

From the above discussion, the random error in timestamps of gateways can have a standard deviation of 5 ns or less depending on the technique used to extract the time from GPS satellites and to synchronize the clocks in the gateways. Localization methods proposed in this proposal use only the time stamps to compute the location of the target. While the error in timestamps of the gateways can be less than 5 ns, or even 2.5 ns depending on the time recovery technique used, unfortunately it is not the only source of error.

The accuracy of the "time stamp" in a gateway has two components. One component is the accuracy of the clock that has been loosely synchronized to the clock in a GPS satellite. The other component is error in recognizing the point in time the message has been received. The accuracy of the latter depends on the bandwidth and duration of the message as well as the signal-to-noise-ratio (SNR) at the input to the receiver. The error in the latter will not be investigated in this Ph. D. research. However, whatever algorithm is proposed, it has to work with the accuracy provided by any of the commercially available gateways, which could be a standard deviation of 25ns or even larger [10, 11].

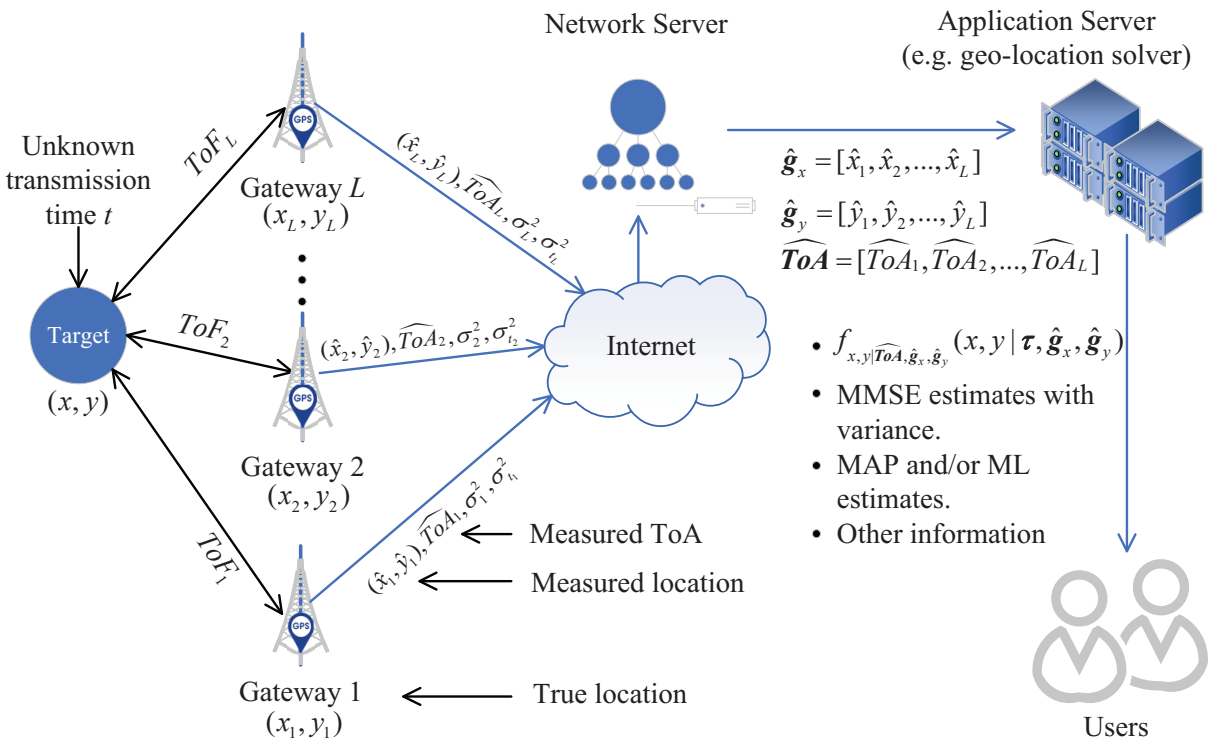

Figure 1.4: Geo-localization system with L gateways.

1.3 Scope of the Research

As mentioned, the objective of this research is to find algorithms that run in the application server with two goals. The primary goal is to determine the probability that the target, i.e., the class A device, is either inside or outside a specified region. The secondary goal is to estimate the location of the target with minimum mean squared error. The inputs to the algorithm are the geo-locations of the multiple gateways that receive the message and the "time stamps" that are inserted in the messages by the gateways. Of course, the time a gateway receives the message will depend on its distance from the sensor. This is illustrated in Figure 1.4, where the gateways and the target are assumed to lie on a planar surface and the surface location of the gateways indicated by (x_ℓ, y_ℓ), $\ell = 1, 2, \ldots, L$, and the surface location of the target is indicated by (x, y).

The time stamp inserted in the message by Gateway ℓ will be the measured time the message arrives at Gateway ℓ. This time is referred to as the estimated Time-of-Arrival

(ToA) and is given by

$$\widehat{\mathsf{ToA}}_\ell = t + \mathsf{ToF}_\ell + \Delta T_\ell, \tag{1.2}$$

where ΔT_ℓ is the measurement error in the "time stamp" of Gateway ℓ and ToA_ℓ is the true time-of-arrival given by

$$\widehat{\mathsf{ToA}}_\ell = t + \mathsf{ToF}_\ell + \Delta T_\ell, \tag{1.3}$$

where ToF_ℓ is the time of flight from the target to Gateway ℓ and t is the exact time of transmission of the message. The errors, ΔT_ℓ, $\ell = 1, 2, \ldots, L$, in (1.2) are caused in part by loose time synchronization to the GPS satellites and in part by Johnson noise at the input to the low noise amplifier. Therefore, $\Delta T_\ell,\ \ell = 1, 2, \ldots, L$ are random errors, independent of each other and reasonably modeled as Gaussian.

The times of flight from the target to the gateways depend on the lengths of the propagation paths. In a line-of-sight environment, which is an environment where there is no obstacle in the straight line between the antennae of the target and the gateway, the time-of-flight is the speed of light times the Euclidean distance between two antennae. This scenario is called a line-of-sight (LoS) scenario. For a LoS path the time of flight from the target to Gateway ℓ is given by

$$\mathsf{ToF}_\ell = \frac{\sqrt{(x_\ell - x)^2 + (y_\ell - y)^2}}{c}. \tag{1.4}$$

In urban environments often the LoS path is blocked and the propagation from the target to the gateway is indirect via a reflection from a building or other structure. Such an environment is referred to as a non-line-of-sight (NLoS). A NLoS scenario is illustrated in Figure 1.5.

The presence of NLoS paths makes it impossible to use multi-lateration[‡] to find the location of the target. In these urban environments the application server has to sort out which gateways are linked to the target by NLoS paths and either discard them or make some sort of correction to minimize the effect of the excessively long propagation paths.

[‡]Multi-lateration is the localization of a target using the times-of-arrival of a burst transmission from the target at a system of spatially separated sensors.

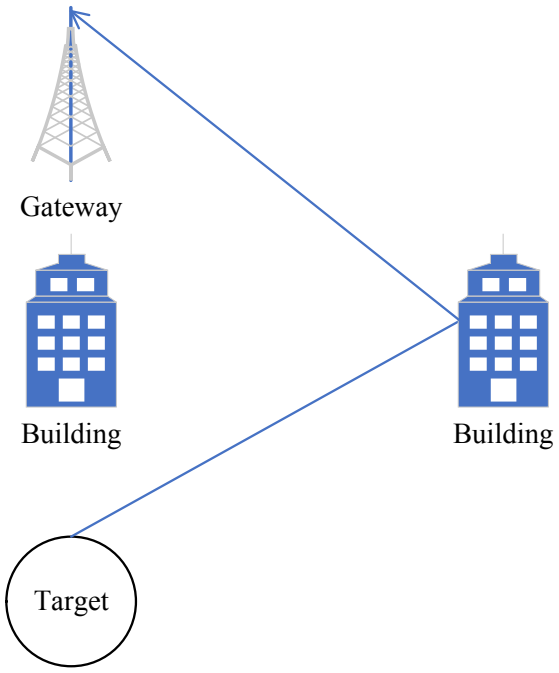

Figure 1.5: Non-line-of-sight path.

The objective is to *localize* a target with a solver, which is a program that runs on a computer, connected to the application server that uses the time stamp information inserted by several gateways into the single message they relay from a Class A end device to the network server. Here the word "localize" is used to mean both find the probability the target is either inside or outside a specified areas and find a point estimate that minimizes the mean squared error.

It should be pointed out that, beside the time of arrival, the received-signal strength (RSS) is usually obtained at the gateway as well. The set of RSS measurements can also be used for the purpose of localization. While RSS-based localization algorithms [12–17] are generally simpler than ToA-based algorithms, they are very sensitive to the path-loss model used to convert RSS to distance. ToA-based localization algorithms [18–26] are generally much more accurate than RSS-based algorithms, but they require precise clock synchronization. It is possible to combine RSS-based and ToA-based algorithms to achieve better performance [27–29]. It will be explained how RSS information can be used to improve performance of the techniques presented in this thesis.

LoRaWAN is a relatively new technology, but geo-localization is not a new topic. Cer-

tainly GPS satellites have been used for localization for years. A great deal of research has been done on *localizing* things like spacecraft orbiting the earth, the buoys tethered to a fisherman's lobster traps, adversarial submarines deep in the ocean, assets in transit and ear-tagged wildlife. While the localization of LoRaWAN compatible sensors is similar to localization of other objects, it is sufficiently different to warrant independent investigation. The objective of this study is to efficiently and accurately localize LoRa end devices (i.e., LoRa sensors) that are referred to as the targets to be localized.

A major part of this research is to investigate algorithms used in other applications and, where possible, relate those algorithms to the geo-localization of LoRa targets. Unfortunately, the algorithms described in the literature use a variety of notational systems that consist of systems of symbols, conventions (e.g. use a bar over a variable to indicate a time average) and unique terminology. Estimation algorithms are communicated by engineers and scientists mathematically using probability and statistics, which is an area of mathematics that has an elaborate notational system that includes a large number of symbols, large set of conventions and many mathematical terms. There is no universal system of symbols and conventions for probability and statistics. The notational systems, which are quite complex, vary from application area to application area and also from country to country. However, probability theory underpins all of the different notational systems, so it is certainly possible to translate one system to another. Unfortunately, the notational systems are quite complicated making it very difficult for engineers and scientists to think in terms of more than one system. Real-time symbol-by-symbol translation while reading a paper diffuses the arguments and leads to misunderstanding.

In order to properly describe work done by others, considerable effort will be spent on both explaining and translating symbols and notational conventions found in the literature. Great care will be taken in defining the problem and tying the symbols and conventions to the physics that underpins the problem. Unfortunately, paying great attention to details will not help those who are familiar with the system of symbols presented in this document. Such readers may find the next Chapter tedious, but for them, with this warning of tediousness, the next Chapter should be a quick read.

2. The Problem, Existing Solutions and the Symbolism

2.1 The Physics of the Problem

In this chapter, the problem is simplified by assuming that the target and the gateways all lie in the same plane. It is referred to as the 2-dimensional localization problem. Later, the problem will be expanded to 3-dimensions. The spatial geometry of the two-dimensional problem is illustrated in Figure 2.1.

It is pointed out that the gateways are typically mounted on towers or the top of tall buildings and will, in most cases, be higher than the target. However, it will be shown later including the third dimension will in general not provide useful height information as the standard deviation in the height estimator will often exceed the roughness of the terrain in most applications.

2.2 Probabilistic Model of the Physics for an LoS Environment

Probabilistic assessment, by its nature, involves statistical analysis that includes random variables and their associated density functions. The foundation for almost all probabilistic problems tackled in scientific papers involves "probabilistic experiments" (page 23 in [30]). Papers in the literature that develop algorithms for the estimator of a random variable usually let the reader imagine the probabilistic experiment. It is believed by the author that a very clear understanding of the probabilistic experiment is necessary to fully comprehend an algorithm that makes a probabilistic assessment or estimates a parameter. For that reason the model for determining the probability that the geo-location of a LoRaWAN end-device is

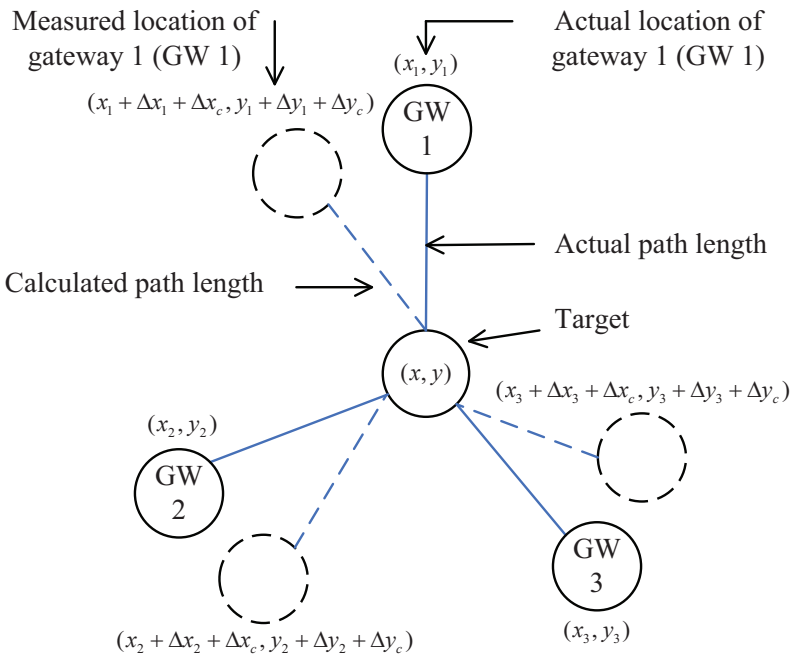

Figure 2.1: Spacial geometry of the problem.

within a specified area is constructed in detail starting with the definition of the probabilistic experiment. While, for the case at hand, the probabilistic experiment could have been left for the reader to imagine, an explicit definition is very helpful to understand the algorithm and the results it produces.

A probabilistic experiment is performed by drawing a single element from a set $\mathcal{S}$ [30]. The element that is drawn is referred to as the outcome of the experiment. The set $\mathcal{S}$ is referred to as the sample space and, by definition, contains all possible outcomes of the experiment. An experiment also requires a probability function. The probability function, denoted $P(\cdot)$ [30], maps each possible event into a "probability of occurrence", where an event is a subset of $\mathcal{S}$. For the problem at hand, the sample space has a countable infinity of elements and is denoted $\mathcal{S} = \{\zeta_1, \zeta_2, \dots\}$. If after performing the experiment the outcome is ζ_i, then any event containing outcome ζ_i is said to have occurred. For example, if event G is the subset $\{\zeta_1, \zeta_2, \zeta_3\}$ and the outcome of the experiment is either ζ_1, ζ_2, or ζ_3 then event G occurs. Otherwise event G does not occur.

Each element of $\mathcal{S}$ is an I-tuple consisting of the geo-location of the target, the time the

message is transmitted, the measured geo-locations of the L gateways in range of the target and the measured times the message arrives at the gateways as well as the errors on the measurements.

The probability function is completely defined if the probability of occurrence of each elementary event $\{\zeta_i\}$, $i = 1, 2, \ldots$, is defined. For the analysis carried out in this thesis, the probability function will be defined implicitly through the definition of several random variables and their density functions.

Random Variables (RVs) are, by definition, functions with domain $\mathcal{S}$ and counter-domain the real line. Since the domain $\mathcal{S}$ consists of outcomes $\zeta_1,\ \zeta_2, \ldots$ the generic argument of every RV will be symbolized with $\boldsymbol{\zeta}$. For example, $x(\boldsymbol{\zeta})$ represents the random variable "x", with argument $\boldsymbol{\zeta}$ being a variable that could be any one of ζ_1, ζ_2, ...

Prior to performing the experiment a random variable, $x(\boldsymbol{\zeta})$, could be any of the values $x(\zeta_1)$, $x(\zeta_2)$, ... After the experiment is performed and the outcome is known, say it is ζ_i, then random variable $x(\boldsymbol{\zeta})$ becomes $x(\zeta_i)$ and has just one value. Note the value that a random variable takes on depends entirely on the outcome. If the outcome is ζ_2, then $x(\boldsymbol{\zeta})$ has the value $x(\boldsymbol{\zeta})\big|_{\boldsymbol{\zeta}=\zeta_2} = x(\zeta_2)$. Prior to performing the experiment $x(\boldsymbol{\zeta})$ is referred to as a random variable. After the experiment is performed the value that random variable $x(\boldsymbol{\zeta})$ takes on is denoted $x(\zeta)$, which differs in notation in that ζ is not bold. That is $x(\zeta)$ is an observation or an instance of random variable $x(\boldsymbol{\zeta})$. The symbol $\boldsymbol{\zeta}$ can be viewed as a variable representing all possible outcomes and ζ is the outcome after the experiment is performed. It is both easy and unambiguous to refer to $x(\zeta)$ when the experiment has been performed and the outcome is known. For example, if the outcome is ζ_2 then the post experiment value is $x(\zeta_2)$. However, there are many occasions where a reference is made to the post experiment value without knowing the outcome. That is, after the experiment has been performed $x(\boldsymbol{\zeta})$ is observed to be a value, say 3.9, but the specific outcome is not known. In such cases, the observation is written $x(\zeta)$, where ζ is not bold.

In order to remove ambiguity "ζ" will be in bold if $x(\boldsymbol{\zeta})$ is used in the context of a random variable, i.e., pre-experiment, and "ζ" will be in regular font if $x(\zeta)$ is either an observation

or a hypothesized outcome. Stated simply, $x(\boldsymbol{\zeta})$ denotes a random variable and $x(\zeta)$ denotes an observation of the random variable.

RVs are completely characterized by their Probability Density Functions (pdf). The probability density function for RV $x(\boldsymbol{\zeta})$ is denoted $f_x(a)$, where the subscript x indicates the density function is associated with RV $x(\boldsymbol{\zeta})$ and the argument "a" is a real number. The density function $f_x(a)$ could have been defined as: "The function $f_x(a)$ whose area between $x = -\infty$ and $x = a$ is the probability that the outcome ζ belongs to the set of outcomes that make $x(\zeta)$ less than a.". In mathematical terms $f_x(a)$ is defined as any function that satisfies

$$P(\{\zeta : x(\boldsymbol{\zeta}) \leq a\}) = \int_{-\infty}^{a} f_x(x)dx, \tag{2.1}$$

where the set $\{\zeta : x(\boldsymbol{\zeta}) \leq a\}$ is the set of all ζ_i's that satisfy $x(\zeta_i) \leq a$. Each random variable has a specific density function, but the converse is not true.

In experiments where several random variables are defined their all of the probabilistic information is contained in joint probability density function (pdf). The individual pdfs, which are called marginal pdfs, do not contain all of the probabilistic information. The joint density function for three random variables $x(\zeta)$, $y(\zeta)$ and $t(\zeta)$, which is denoted $f_{x,y,t}(x, y, t)$, is defined as any function $f_{x,y,t}(x, y, t)$ satisfying

$$P(\{\zeta : x(\boldsymbol{\zeta}) \leq x_1,\ y(\boldsymbol{\zeta}) \leq y_1,\ t(\boldsymbol{\zeta}) \leq t_1\}) = \int_{-\infty}^{t_1}\int_{-\infty}^{y_1}\int_{-\infty}^{x_1} f_{x,y,t}(x, y, t)dxdydt, \tag{2.2}$$

where the set $\{\zeta : x(\boldsymbol{\zeta}) \leq x_1,\ y(\boldsymbol{\zeta}) \leq y_1,\ t(\boldsymbol{\zeta}) \leq t_1\}$ is the set of all ζ_i's that simultaneously satisfy the three inequalities $x(\boldsymbol{\zeta}) \leq x_1$, $y(\boldsymbol{\zeta}) \leq y_1$, and $t(\boldsymbol{\zeta}) \leq t_1$. The subscripts in the joint density function $f_{x,y,t}(x, y, t)$ are vital as they associate the argument with the random variable. For example, $f_{x,y,t}(1, 2, 3)$ indicates the arguments associated with RVs $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$, and $t(\boldsymbol{\zeta})$ are 1, 2 and 3, respectively.

There are $5L + 5$ fundamental random variables defined in this chapter. None of the $5L + 5$ fundamental of random variables is observable, which means none can be measured

Table 2.1: Random variables used in the analysis for a system with L gateways.

Random Variable	**Variance**	**Description**
$t(\boldsymbol{\zeta})$	$\approx \infty$	The time at which the target transmits the message
$x(\boldsymbol{\zeta})$	$\approx \infty$	The x-coordinate of the target
$y(\boldsymbol{\zeta})$	$\approx \infty$	The y-coordinate of the target
$x_\ell(\boldsymbol{\zeta})$	$\approx \infty$	The x-coordinate of Gateway ℓ
$y_\ell(\boldsymbol{\zeta})$	$\approx \infty$	The y-coordinate of Gateway ℓ
$\Delta x_c(\boldsymbol{\zeta})$	N/A	The component of error in the x-coordinate common to the location of all gateways
$\Delta y_c(\boldsymbol{\zeta})$	N/A	The component of error in the y-coordinate common to the location of all gateways
$\Delta x_\ell(\boldsymbol{\zeta})$	$\sigma^2_{x_\ell} = \sigma^2_\ell$	The component of the random error in measurement of the x-coordinate of the location of Gateway ℓ
$\Delta y_\ell(\boldsymbol{\zeta})$	$\sigma^2_{y_\ell} = \sigma^2_\ell$	The component of the random error in measurement of the y-coordinate of the location of Gateway ℓ
$\Delta T_\ell(\boldsymbol{\zeta})$	$\sigma^2_{t_\ell}$	The error in measuring the time the message arrives at Gateway ℓ

without error. These fundamental random variables are listed in Table 2.1 and described below:

$t(\boldsymbol{\zeta})$: This is the time that the target transmits the message. $t(\boldsymbol{\zeta})$ is independent of all other RVs. Its probability density function is assumed to be Gaussian with a mean of zero and a variance that tends to ∞. It could have been defined as uniformly distributed in the interval $(-a, a)$ where a is very large. Such a distribution will be referred to as widely uniform. In Bayesian terminology, density functions based on the physics that underpins the application or other pre-experiment knowledge is referred to as the *a priori* pdf. Therefore, this zero-mean large-variance Gaussian pdf (i.e., this widely uniform pdf) is the *a priori* pdf for $t(\zeta)$.

This density function is referred to as the *a prior* pdf.

$x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$: These two RVs are the x and y coordinates of the target. They may or may not be independent of each other. They have a joint density, $f_{x,y}(x, y)$, that is known from the physics of the application. Often the physics does not provide information about $f_{x,y}(x, y)$ in which case it is taken to be zero-mean, which is equivalent to being

uniformly distributed over the region $-a < x < a$ and $-a < y < a$ for very large a, uncorrelated Gaussian with marginal variances that tend to ∞.

In Baysian terminology this probability density function is referred to as the *a priori* pdf.

$x_\ell(\boldsymbol{\zeta})$, $y_\ell(\boldsymbol{\zeta})$: These two RVs are the x and y coordinates of Gateway ℓ. They are assumed to be independent of each other and all other RVs and have zero-mean Gaussian distributions with variances that tend to ∞. Since there are L gateways, there are $2L$ of this type of RVs. Obviously, measured values of these random variables are used in the localization algorithm.

$\Delta x_c(\boldsymbol{\zeta})$, $\Delta y_c(\boldsymbol{\zeta})$: These two RVs represent the component of error common to the location of all gateways. The error is due to the variation in the orbits of the GPS satellites (probably caused by the moon and the sun shifting the center of the gravity so the plane of the satellites orbit does not go through the center of the earth). They are assumed to be independent of all other RVs. However, no assumption is made about the shape of their joint density and whether or not they are independent of each other.

$\Delta x_\ell(\boldsymbol{\zeta})$, $\Delta y_\ell(\boldsymbol{\zeta})$: These two RVs represent the random errors in the measurements of the x and y coordinates of Gateway ℓ. Should these coordinates be measured using the GPS, the variance of this error will depend on the strength of the received satellite signal. Since there are L gateways, there are $2L$ of this type of RVs. They are assumed to be independent of all other RVs as well as independent of each other. They are also assumed to be zero-mean Gaussian RVs with identical variances denoted $\sigma^2_{x_\ell} = \sigma^2_{y_\ell} = \sigma^2_\ell$. These variances are assumed to be computed by the gateway and their values placed inside the messages sent to the network server.

$\Delta T_\ell(\boldsymbol{\zeta})$: This is the error in measuring the time a message arrives at Gateway ℓ. These RVs represent the sum of two types of error:

i. The random error introduced in the process that attempts to synchronize the gateway's clock to the GPS satellites' clocks, and

ii. The error due to Johnson noise in the receiver blurring the point in time a message is received.

There are L such random variables. They are assumed to be mutually independent and also independent of all other fundamental random variables. They are assumed to be zero-mean Gaussian RVs with possibly different variances, denoted $\sigma_{t_\ell}^2$. The variances of these RVs are calculated in the gateways at the time a message is received and inserted into the message sent to the network server.

The errors represented by $\Delta x_c(\boldsymbol{\zeta})$, $\Delta y_c(\boldsymbol{\zeta})$, $\Delta x_\ell(\boldsymbol{\zeta})$ and $\Delta y_\ell(\boldsymbol{\zeta})$ are in part due to limitations of the GPS satellites and in part due to Johnson noise at the input to the receiver. The physics governing the motion of the GPS satellites, most notably the diurnal variations in the satellites orbits, and the physics that underpins Johnson noise (i.e., thermal noise) together with the circuits/algorithms that generate the measurements determine the joint density function of the $5L + 5$ fundamental random variables listed in Table 2.1.

The majority of the random error in the measurements of the gateways' locations and the times-of-arrival is caused by the Johnson noise at the inputs of the low-noise amplifiers in the gateways' receivers. Since Johnson noise has a Gaussian amplitude distribution and any linear operation on a signal with a Gaussian amplitude distribution produces a signal with a Gaussian amplitude distribution, it is expected that the errors caused by the Johnson noise will have a Gaussian amplitude distribution.

The joint distribution function for the random errors is a little more difficult to predict. There are L gateways, but the signals from the target and the GPS satellites occupy different frequency bands therefore there are $2L$ sources of Johnson noise affecting $3L$ measurements. Even so, it is believed that the errors in the $3L$ random measurements are independent.

The errors common to the gateways' locations, i.e., $\Delta x_c(\boldsymbol{\zeta})$ and $\Delta y_c(\boldsymbol{\zeta})$ are unrelated to the Johnson noise so they are independent of the $3L$ random errors. No restrictions have been placed on the joint density of the common error. The analysis to follow will be careful to allow any joint distribution for $\Delta x_c(\boldsymbol{\zeta})$ and $\Delta y_c(\boldsymbol{\zeta})$. To be specific, the analysis to follow will consider the effect of $\Delta x_c(\boldsymbol{\zeta})$ and $\Delta y_c(\boldsymbol{\zeta})$ on the estimate of the target's location separately.

This leaves the *a priori* distributions for the x-y coordinates of the target, the *a priori* distribution of the time that the target transmits the message and the *a priori* distribution for the x-y coordinates of the L gateways. In many estimation problems some quantities are completely unknown. Often these unknown quantities are modeled with an independent uniform distribution spanning from $-\infty$ to ∞. Equivalently, these unknown quantities can be modeled with a zero-mean Gaussian distribution with a variance that tends to infinity, which will be done in this thesis. The coordinates of the target, the time of transmission, as well as the $2L$ coordinates of the gateways will be modeled as independent zero-mean Gaussian random variables with variances that tend to infinity. This is done to simplify the mathematics.

Having established $\Delta x_\ell(\boldsymbol{\zeta})$, $\Delta y_\ell(\boldsymbol{\zeta})$, and $\Delta T_\ell(\boldsymbol{\zeta})$ for $\ell = 1, 2, \ldots, L$, as independent zero-mean Gaussian random variables, their distributions are characterized by their variances. The variances for $\Delta x_\ell(\boldsymbol{\zeta})$, $\Delta y_\ell(\boldsymbol{\zeta})$, and $\Delta T_\ell(\boldsymbol{\zeta})$ are denoted $\sigma^2_{x_\ell}$, $\sigma^2_{y_\ell}$, and $\sigma^2_{t_\ell}$, respectively, where it is understood that $\sigma^2_{x_\ell} = \sigma^2_{y_\ell} = \sigma^2_\ell$.

Random variables $\Delta x_c(\boldsymbol{\zeta})$ and $\Delta y_c(\boldsymbol{\zeta})$ likely do not have Gaussian distributions. The analysis to follow does not use their joint distribution, i.e., the analysis does not use

$$f_{\Delta x_c, \Delta y_c}(\Delta x_c, \Delta y_c).$$

It is important enough to repeat that the *a priori* distributions for unknown random variables $t(\boldsymbol{\zeta})$, $x_\ell(\boldsymbol{\zeta})$ and $y_\ell(\boldsymbol{\zeta})$ are assumed to be independent zero-mean and Gaussian distributed with variances that tend to ∞. The *a priori* joint density for $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ may be known, in which case it could be any valid pdf, or it may be unknown, in which case it could be assumed to be zero-mean uncorrelated Gaussian with marginal variances that tends to be ∞.

Lastly, random variables $\Delta x_c(\boldsymbol{\zeta})$ and $\Delta y_c(\boldsymbol{\zeta})$ are applied to the estimated location of target after the target's location is estimated. Shifting the location of all gateways by $(\Delta x_c(\zeta), \Delta y_c(\zeta))$ prior to estimating the location of the target is equivalent to estimating the location of the target without shifting the locations of the gateways and then shifting the

estimated location of the target by $(\Delta x_c(\zeta), \Delta y_c(\zeta))$. The common error in gateway location must be applied to the post estimation location of the target. This approach is necessary to make the analysis tractable when $f_{\Delta x_c,\, \Delta y_c}(\Delta x_c,\ \Delta y_c)$ is not Gaussian. The effect of the error common to all gateways is not included in any examples in this study.

2.3 Mathematical Model for Localization in an LoS Environment

The approach that will be taken is to consolidate the information into L equations, 1 for each gateway. The L equations must relate the coordinates of the target to the measurements of the times-of-arrival (ToA) and the geo-locations of the gateways. The relationship of interest is established through the time-of-flight (ToF) equations. Each time-of-arrival measurement is a sum of: (i) the time the message is transmitted, (ii) the time of flight and (iii) the error in measuring the arrival time.

The convention that will be used to distinguish measurements from the exact quantity is to symbolize the measurement by placing a "hat" on top of the symbol for the exact value. Since all measured quantities will be functions of the outcome of the experiment, measurements are instances of a random variable. For example, the set of possible measurements of the random variable representing all of the times-of-arrival at Gateway ℓ is symbolized by the random variable $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ (note: $\boldsymbol{\zeta}$ is bold for random variables) and the post experiment measurement value of the time-of-arrival will be symbolized $\widehat{\mathsf{ToA}}_\ell(\zeta)$ (note: ζ is regular font for observations, i.e., instances of random variables), where ζ represents the outcome of the experiment and will be one of ζ_1, ζ_2, ... This implies the random variable representing the set of all of the possible exact times-of-arrival is denoted $\mathsf{ToA}_\ell(\boldsymbol{\zeta})$ and its post experiment value, which is also referred to as the observation, is $\mathsf{ToA}_\ell(\zeta)$.

The L equations for the RV $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ are

$$\begin{aligned}\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta}) &= \mathsf{ToA}_\ell(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta}) \\ &= t(\boldsymbol{\zeta}) + \mathsf{ToF}_\ell(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta})\end{aligned} \tag{2.3}$$

and the L equations for the post-experiment observations are

$$\begin{aligned}\widehat{\mathsf{ToA}}_\ell(\zeta) &= \mathsf{ToA}_\ell(\zeta) + \Delta T_\ell(\zeta) \\ &= t(\zeta) + \mathsf{ToF}_\ell(\zeta) + \Delta T_\ell(\zeta) \end{aligned} \tag{2.4}$$

respectively, where $t(\zeta)$ is the time of transmission, $\Delta T_\ell(\zeta)$ is the measurement error in the time the message arrives at Gateway ℓ, i.e., the error in the time stamp that is inserted in the message by the gateway, and $\mathsf{ToF}_\ell(\zeta)$ is the time-of-flight from the target located at $(x(\zeta), y(\zeta))$ to Gateway ℓ located at $(x_\ell(\zeta), y_\ell(\zeta))$.

The time-of-arrival measurements together with the measurements of the coordinates for the gateways are used to determine the probability the target is within a specified area and to estimate the coordinates of the target. The analysis, which is presented in detail in Chapter 3, provides the joint probability density function (pdf) of the target's coordinates (x, y). In order to contrast the approach taken in this research which is to find the Bayesian *a posteriori* pdf, against point estimation techniques, the next sections review classical and state-of-the-art methods used to find point estimates of the target's location.

It will be made clear that none of the previous work has provided a minimum mean squared error point estimate and that no literature has been found on determining the probability that the target is in a specified area. That is, a literature search did not turn up any previous work on determining the *a posteriori* pdf for the coordinates of a target from time-of-arrival measurements.

2.4 Classical Methods for the Localization Problem

It must be emphasized that a fairly extensive literature review did not turn up any papers on the probability that the target is in a specified area. However, the literature review did turn up many papers on point estimators for the target's coordinates. It is pointed out the none of these papers on point estimators derived the minimum mean squared error estimator.

In the literature, the time difference of arrivals (TDoA) is widely used to deal with the unknown transmission time of the target [31]. Specifically, $L - 1$ TDoA values are obtained

by subtracting one of the measured ToA values from the other $L-1$ values so as to remove the unknown transmission time. For example, using the first ToA value as a reference, one obtains

$$\widehat{\mathsf{TDoA}}_\ell(\boldsymbol{\zeta}) = \widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta}) - \widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}), \quad \ell = 2, 3, \ldots, L, \tag{2.5}$$

The observed TDoA values together with the coordinates of the gateways are used to estimate the (x, y) coordinates of the target. First, $\widehat{\mathsf{TDoA}}_\ell(\zeta)$ is related to the difference between the ranges from the target to Gateways ℓ and 1 by substituting (2.3) for $\widehat{\mathsf{ToA}}_\ell(\zeta)$ in (2.5) and multiplying both sides of the result by the speed of light, c. The measured range difference between Gateway ℓ and Gateway 1, denoted $\widehat{d}_{\ell,1}(\zeta)$, is given by

$$\begin{aligned}
\widehat{d}_{\ell,1}(\zeta) &= c\widehat{\mathsf{TDoA}}_\ell(\zeta) \\
&= c\big(\mathsf{ToF}_\ell(\zeta) + \Delta T_\ell(\zeta)\big) - c\big(\mathsf{ToF}_1(\zeta) + \Delta T_1(\zeta)\big) \\
&= c(\mathsf{ToF}_\ell(\zeta) - \mathsf{ToF}_1(\zeta)) + c\big(\Delta T_\ell(\zeta) - \Delta T_1(\zeta)\big) \\
&= c(\sqrt{(x_\ell - x(\zeta))^2 + (y_\ell - y(\zeta))^2} - \sqrt{(x_\ell - x(\zeta))^2 + (y_\ell - y(\zeta))^2}) \\
&\qquad + c\big(\Delta T_\ell(\zeta) - \Delta T_1(\zeta)\big), \quad \ell = 2, 3, \ldots, L,
\end{aligned} \tag{2.6}$$

where $\mathsf{ToF}_\ell(\zeta) = \sqrt{(x_\ell - x(\zeta))^2 + (y_\ell - y(\zeta))^2}$. Since $\widehat{d}_{\ell,1}(\zeta)$ is a measurement it is an observation of random variable

$$\begin{aligned}
\widehat{d}_{\ell,1}(\boldsymbol{\zeta}) &= c(\sqrt{(x_\ell - x(\boldsymbol{\zeta}))^2 + (y_\ell - y(\boldsymbol{\zeta}))^2} - \sqrt{(x_\ell - x(\boldsymbol{\zeta}))^2 + (y_\ell - y(\boldsymbol{\zeta}))^2}) \\
&\qquad + c\big(\Delta T_\ell(\boldsymbol{\zeta}) - \Delta T_1(\boldsymbol{\zeta})\big), \quad \ell = 2, 3, \ldots, L,
\end{aligned} \tag{2.7}$$

Then, if $x(\zeta)$ and $y(\zeta)$ are given, say to be x and y, then $\widehat{d}_{\ell,1}(\boldsymbol{\zeta})$ has a Gaussian distribution since $\Delta T_\ell(\boldsymbol{\zeta})$ has a Gaussian distribution.

In order to reduce notional clutter let $\widehat{d}_{\ell,1} = \widehat{d}_{\ell,1}(\zeta)$, the vector $\widehat{\boldsymbol{d}}$ be $\widehat{\boldsymbol{d}} = (\widehat{d}_{2,1}, \ldots, \widehat{d}_{L,1})$ and $\widehat{\boldsymbol{d}}(\boldsymbol{\zeta})$ be the vector of random variables $\widehat{\boldsymbol{d}}(\boldsymbol{\zeta}) = (\widehat{d}_{2,1}(\boldsymbol{\zeta}), \ldots, \widehat{d}_{L,1}(\boldsymbol{\zeta}))$ since $\widehat{\boldsymbol{d}}_{\ell,1}(\zeta)$ is Gaussian if $x(\zeta)$ and $y(\zeta)$ are given. Then, the joint probability density function of $\widehat{d}_{2,1}(\boldsymbol{\zeta}), \ldots, \widehat{d}_{L,1}(\boldsymbol{\zeta})$

conditioned on the target location is a multivariate Gaussian density. It is given by [31]

$$f_{\widehat{\boldsymbol{d}}\big|x,y}(\widehat{\boldsymbol{d}}\big|x,y) = \frac{1}{\sqrt{(2\pi)^{L-1}|\boldsymbol{\Phi}|}} \exp\left(-\frac{1}{2}(\textsf{range\_diff})^\top \boldsymbol{\Phi}^{-1}(\textsf{range\_diff})\right), \tag{2.8}$$

where $\boldsymbol{\Phi}$ is the covariance matrix of RVs $\widehat{d}_{2,1}(\boldsymbol{\zeta}), \ldots, \widehat{d}_{L,1}(\boldsymbol{\zeta})$, given by [32, 33]

$$\boldsymbol{\Phi} = \begin{bmatrix} c^2(\sigma_{t_2}^2 + \sigma_{t_1}^2) & c^2(\sigma_{t_1}^2) & \ldots & c^2(\sigma_{t_1}^2) \\ c^2(\sigma_{t_1}^2) & c^2(\sigma_{t_3}^2 + \sigma_{t_1}^2) & \ldots & c^2(\sigma_{t_1}^2) \\ \vdots & \vdots & \ldots & \vdots \\ c^2(\sigma_{t_1}^2) & c^2(\sigma_{t_1}^2) & \ldots & c^2(\sigma_{t_L}^2 + \sigma_{t_1}^2) \end{bmatrix}.$$

and **range_diff** is the vector

$$\begin{bmatrix} \widehat{d}_{2,1} - \sqrt{(x_2 - x)^2 + (y_2 - y)^2} + \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \\ \widehat{d}_{3,1} - \sqrt{(x_3 - x)^2 + (y_3 - y)^2} + \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \\ \vdots \\ \widehat{d}_{L,1} - \sqrt{(x_L - x)^2 + (y_L - y)^2} + \sqrt{(x_1 - x)^2 + (y_1 - y)^2} \end{bmatrix}.$$

The target's location is often estimated with the (x, y) coordinates which are in some sense most consistent with the observation $\widehat{\boldsymbol{d}}$. Within this class, different criteria have been introduced in the literature to measure the consistency. These include (i) position that minimizes the sum of squares of the errors in $L - 1$ range differences, and (ii) position that is most likely to occur given the measured range differences. Furthermore, estimators that adopt these criteria can be classified further as either non-linear or linear [31]. They are discussed in the following subsections.

2.4.1 Non-Linear Estimators

The non-linear least squares (NLS) estimator minimizes the sum of the squares of the error in the measured range differences. Specifically, the NLS estimator minimizes the cost function:

$$\boldsymbol{J}_{\text{NLS}} = (\textsf{range\_diff})^\top (\textsf{range\_diff}). \tag{2.9}$$

On the other hand, the *maximum-likelihood* (ML) estimator finds a location that maximizes the joint pdf in (2.8). For each hypothesis of the target location (x, y), the joint pdf in (2.8) essentially measures the likelihood of measurements $\widehat{d}_{\ell,1}(\zeta)$, hence the name *maximum-likelihood* estimator. Since the $\ln(\cdot)$ function is monotonic increasing, such an estimator is equivalent to the one that maximizes the log of the likelihood function of the measurements $\widehat{d}_{\ell,1}(\zeta)$, which is:

$$\ln(f_{\widehat{\boldsymbol{d}}|x,y}(\widehat{\boldsymbol{d}} \mid x, y) = \ln\left(\frac{1}{\sqrt{(2\pi)^{L-1}|\boldsymbol{\Phi}|}}\right) - \frac{1}{2}(\textsf{range\_diff})^{\top}\boldsymbol{\Phi}^{-1}(\textsf{range\_diff}). \tag{2.10}$$

Since the first term on the right side of (2.10) is a constant, maximizing (2.10) is equivalent to minimizing

$$\boldsymbol{J}_{\mathrm{ML}} = (\textsf{range\_diff})^{\top}\boldsymbol{\Phi}^{-1}(\textsf{range\_diff}). \tag{2.11}$$

Whether the NLS or ML estimator is used, the problem becomes finding (x, y) to minimize J_{NLS} or J_{ML}, respectively. Since a closed-form solution is very difficult, if not impossible, to obtain, the solution is typically found with a computer search that could be a global grid search or a steered iterative search. There are a variety of algorithms available for conducting iterative searches, including genetic algorithms, particle swam optimization, Newton-Raphson, Gauss-Newton, and steepest descent [31].

Among these iterative search methods, the steepest descent method is perhaps the simplest and most stable. This method solves the problem of minimizing (2.9) or (2.10) iteratively as follows:

$$\begin{aligned} \widetilde{\boldsymbol{x}}^{(n+1)} &= \widetilde{\boldsymbol{x}}^{(n)} - \lambda \nabla \boldsymbol{J}_{\mathrm{NLS}}^{(n)} \\ \text{and} \quad \widetilde{\boldsymbol{x}}^{(n+1)} &= \widetilde{\boldsymbol{x}}^{(n)} - \lambda \nabla \boldsymbol{J}_{\mathrm{ML}}^{(n)} \end{aligned} \tag{2.12}$$

where $\widetilde{\boldsymbol{x}}^{(n)} = (\widetilde{x}^{(n)}, \widetilde{y}^{(n)})$ is the solution at the nth iteration, $\nabla \boldsymbol{J}_{\mathrm{NLS}}^{(n)}$ and $\nabla \boldsymbol{J}_{\mathrm{ML}}^{(n)}$ are the gradient vectors of $\boldsymbol{J}_{\mathrm{NLS}}$ and $\boldsymbol{J}_{\mathrm{ML}}$ evaluated at $\boldsymbol{x} = \widetilde{\boldsymbol{x}}^{(n)}$, respectively, and λ is a positive constant that controls the convergence rate (it is usually set as small as necessary to ensure stability).

The gradient vectors $\nabla \boldsymbol{J}_{\mathrm{NLS}}^{(n)}$ and $\nabla \boldsymbol{J}_{\mathrm{ML}}^{(n)}$ are obtained by taking the partial derivatives of $\boldsymbol{J}_{\mathrm{NLS}}$ and $\boldsymbol{J}_{\mathrm{ML}}$, respectively, with respect to the target's coordinates x and y and evaluating the results at $\boldsymbol{x} = \widetilde{\boldsymbol{x}}^{(n)}$. In particular, $\nabla \boldsymbol{J}_{\mathrm{NLS}}^{(n)}$ and $\nabla \boldsymbol{J}_{\mathrm{ML}}^{(n)}$ are given by

$$\begin{aligned}
\nabla \boldsymbol{J}_{\mathrm{NLS}}^{(n)} &= 2 \begin{bmatrix} \sum_{\ell=2}^{L} \left(\widehat{\boldsymbol{d}}_{\ell,1} - \widetilde{d}_{\ell,1}^{(n)} \right) \left(-\frac{\widetilde{x}^{(n)} - x_\ell}{\widetilde{r}_\ell^{(n)}} + \frac{\widetilde{x}^{(n)} - x_1}{\widetilde{r}_1^{(n)}} \right) \\ \sum_{\ell=2}^{L} \left(\widehat{\boldsymbol{d}}_{\ell,1} - \widetilde{d}_{\ell,1}^{(n)} \right) \left(-\frac{\widetilde{y}^{(n)} - y_\ell}{\widetilde{r}_\ell^{(n)}} + \frac{\widetilde{y}^{(n)} - y_1}{\widetilde{r}_1^{(n)}} \right) \end{bmatrix} \text{ and} \\
\nabla \boldsymbol{J}_{\mathrm{ML}}^{(n)} &= \begin{bmatrix} \sum_{\ell=2}^{L} \dfrac{2\partial R_\ell}{\partial x} R_\ell b_{\ell-1,\ell-1} \\ \sum_{\ell=2}^{L} \dfrac{2\partial R_\ell}{\partial y} R_\ell b_{\ell-1,\ell-1} \end{bmatrix} + \begin{bmatrix} \sum_{i=2}^{L} \sum_{j=2, j\neq i}^{L} \dfrac{\partial R_i}{\partial x} R_j (b_{i-1,j-1} + b_{j-1,i-1}) \\ \sum_{i=2}^{L} \sum_{j=2, j\neq i}^{L} \dfrac{\partial R_i}{\partial y} R_j (b_{i-1,j-1} + b_{j-1,i-1}) \end{bmatrix}.
\end{aligned} \tag{2.13}$$

where $\widetilde{d}_{i,j}^{(n)} = \widetilde{r}_i^{(n)} - \widetilde{r}_j^{(n)}$, $\widetilde{r}_i^{(n)} = \sqrt{(\widetilde{x}^{(n)} - x_i)^2 + (\widetilde{y}^{(n)} - y_i)^2}$,

$$\frac{\partial R_\ell^{(n)}}{\partial x} = -\frac{\widetilde{x}^{(n)} - x_\ell}{\widetilde{r}_\ell^{(n)}} + \frac{\widetilde{x}^{(n)} - x_1}{\widetilde{r}_1^{(n)}} \text{ and} \qquad \frac{\partial R_\ell^{(n)}}{\partial y} = -\frac{\widetilde{y}^{(n)} - y_\ell}{\widetilde{r}_\ell^{(n)}} + \frac{\widetilde{y}^{(n)} - y_1}{\widetilde{r}_1^{(n)}}.$$

2.4.2 Linear Estimators

Due to the high complexity of non-linear estimators, linear estimators have been developed by converting the non-linear range equation $\boldsymbol{\Lambda} = \boldsymbol{d} + \boldsymbol{n}$ to a set of equations that are linear in the target's location $\boldsymbol{x} = (x, y)^\top$. Specifically, the transformed range equation is

$$\boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{q}, \tag{2.14}$$

where $\boldsymbol{b}$ and $\boldsymbol{A}$ are known, while $\boldsymbol{q}$ is a noise vector transformed from $\boldsymbol{n}$.

Some linear estimators such as linear least squares (LLS) [34, 35], weighted linear least squares (WLLS) [36–39], or subspace [40–43] generally have a lower accuracy than the non-linear estimators, but they have lower computation complexity.

2.4.3 State-of-the-Art Methods

Recently, convex optimization has been used to solve the point-estimate localization problem in wireless sensor networks. Methods developed based on this technique can be

classified as either second-order cone programming (SOCP) or semidefinite programming (SDP) [44]. The SDP programming technique is considered as a generalized form of SOCP when the SOCP constraints are written as a linear matrix inequality.

The work in [45] proposed a SDP method to solve the localization problem under the ideal scenario that the distances from the target to gateways do not have any errors. The authors in [46] extended the problem in [45] to allow errors in distance measurements. The methods in [24, 44] can provide estimates of the target's location when the transmission time is unknown and the measurements of gateways' locations contain errors. Several works [21, 47, 48] have recently solved the ToA-based localization problem in NLoS environments without knowing the transmission time.

As an example, consider solving the localization problem based on the ToA measurements (without knowing the transmission time) by SDP presented in [47]. Skipping all the details, the SDP problem posed by the authors in [47] is as follows:

$$\min_{\boldsymbol{h},\boldsymbol{q},x,y,\boldsymbol{u},z,t,s} \sum_{\ell=1}^{L} m_\ell(c^2\widehat{\mathsf{ToA}}_\ell^{\,2} + c^2 s - 2c^2\widehat{\mathsf{ToA}}_\ell t - h_\ell - q_\ell)^2 + \sum_{\ell=1}^{L}\lambda(q_\ell^2 + s^2) + \gamma\sum_{\ell=1}^{L} u_\ell^2$$

$$\text{subject to:} \quad h_i = \begin{pmatrix} x_\ell \\ y_\ell \\ -1 \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{I_2} & \boldsymbol{0} \\ \boldsymbol{0} & z \end{pmatrix} \begin{pmatrix} x_\ell \\ y_\ell \\ -1 \end{pmatrix};$$

$$\begin{pmatrix} \boldsymbol{I_2} & \boldsymbol{0} \\ \boldsymbol{0} & z \end{pmatrix} \geq 0;$$

$$q_\ell \geq 0;$$

$$\widehat{\mathsf{ToA}}_\ell \geq t;$$

$$(\widehat{\mathsf{ToA}}_\ell^{\,2} - \widehat{\mathsf{ToA}}_\ell t)c^2 + u_\ell \geq h_\ell;$$

$$u_\ell \geq 0, \quad \ell = 1, \ldots, L;$$

$$(\widehat{\mathsf{ToA}}_i - t)c + (\widehat{\mathsf{ToA}}_j - t)c \geq \|(x_i,\ y_i)^\top - (x_j,\ y_j)^\top\|,\ i \neq j; i, j = 1, \ldots, L$$

In the above SDP problem, the objective function is the sum of errors in ToA measurements. The other parameters/variables are as follows:

- t is the unknown transmission time at the target.

- m_ℓ, λ, γ are the optimization variables of the SDP problem, which are found by solving other optimization problems formulated based on a specific assumption of the range difference measurements. In [47], the values of these variables are manually set, which visually achieve a good localization accuracy.

- $\boldsymbol{h} = (h_1, \ldots, h_\ell), \boldsymbol{q} = (q_1, \ldots, q_\ell), \boldsymbol{u} = (u_1, \ldots, u_\ell), z, s$ are variables used to convert the non-linear ToA-based localization problem to a linear convex problem.

The above SDP problem can be solved with optimization tools such as the MATLAB toolbox CVX [21]. In general, the accuracy of the SDP-based solutions strongly depends on many optimization variables and predefined constraints. Besides, the accuracy also depends on the linearization technique used to turn the objective function into a linear polynomial form required by the SDP technique.

2.4.4 Comparison of Existing Point-Estimation Methods

Table 2.2 summarizes different point-estimation methods discussed in previous sections. It is similar to the table on page 35 of [31], but also includes the more recent SDP-based methods. The classical methods in Table 2.2 are presented with a prefix "TDoA" in their names to emphasize that their estimates are the (x, y) coordinates which are in some sense most consistent with the time-differences-of-arrival $\widehat{\boldsymbol{d}}$.

The accuracy of the existing localization methods mostly depends on their objective functions. These objective functions are mainly classified as either linearized or non-linearized. The methods with non-linearized objective functions such as TDoA-ML and TDoA-NLS methods generally have high accuracy when all information of the input data is used to estimate the target's location. The main disadvantage of this class is that the complexity is from moderate to high, especially when the grid or random search is applied to achieve very high accuracy. The methods with linearized objective functions that include SDP, TDoA-LLS and TDoA-WLLS can provide only the point estimate for the target's location with different levels of complexity. Their accuracies vary from low to high depending on the amount of the

Table 2.2: Comparison of different localization methods (point estimation).

Method	**Accuracy**	**Complexity**	**Statistics Required**	**Linearized**
TDoA-ML	Highest accuracy	High if grid or random search is applied	Yes	No
TDoA-NLS	High accuracy, generally lower than the ML method	High if grid or random search is applied	No	No
SDP	Can achieve accuracy of the ML method with constraints	High	No	Yes
TDoA-LLS or subspace	Generally low	Low	No	Yes
TDoA-WLLS	Can achieve accuracy of the ML method with constraints	Low to moderate, may require iterations	Yes	Yes

information from the input data being lost after the linearization process and the preciseness of the constraints to limit the search space.

2.4.5 Approximated Region Estimation Based on the Point Estimation

In contrast to the point estimation paradigm, the region estimation paradigm has yet to be explored. In the following, we summarize the analysis by Foy in [49], which could be extended to obtain a rough Gaussian approximation to the *a posteriori* density of the target's location.

Foy [49] presented an iterative algorithm that finds the LS solution (a point estimate) to the set of non-linear equations and he also provided the approximate covariance matrix for the solution. His technique improves an initial guess for the solution by linearizing (using a first-order Taylor series polynomial) the set of non-linear equations about the guessed solution and then using linear algebra to get the LS solution to the linear equations. The process is repeated using the improved solution as the "guessed solution" until it converges.

Let $(\bar{x}, \bar{y})$ be the final solution obtained by Foy's iterative technique. Then $(\bar{x}, \bar{y})$ serves as the mean of the Gaussian approximation, while the covariance matrix is the 2×2 upper

left block of the following 3×3 matrix:

$$\bar{\mathbf{C}} = c^2 \left(\mathbf{A}^\top \mathbf{C}_{\Delta T}^{-1} \mathbf{A} \right)^{-1}, \tag{2.15}$$

where $\mathbf{C}_{\Delta T} = \text{diag}(\sigma_{t_1}^2, \ldots, \sigma_{t_L}^2)$, and $\mathbf{A}$ is a $L \times 3$ matrix with row ℓ being

$$\left[\frac{\bar{x} - x_\ell}{\sqrt{(\bar{x} - x_\ell)^2 + (\bar{y} - y_\ell)^2}}, \frac{\bar{y} - y_\ell}{\sqrt{(\bar{x} - x_\ell)^2 + (\bar{y} - y_\ell)^2}}, 1 \right] \tag{2.16}$$

3. *A Posteriori* Probability Density of Target's Coordinates and MMSE, MAP Estimators

3.1 Review of Estimation Theory

The roots of estimation theory go back to 1792 when work by Thomas Bayes [50] was published posthumously. The next major change happened in 1933 when mathematician Andrey Kolmogorov published a book titled "Foundation of the Theory of Probability". In that book Kolmogorov set forth 3 axioms (three statements that must be true) and developed probability theory from those axioms. The entirety of probability theory as we know it today could be developed and proven from Kolmogorov's three axioms. Estimation theory, which is a subset of probability theory has not changed significantly in 50 years. Evidence of this lies in two of publications. Van Trees published a book titled "Detection Estimation and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory" [51] in 1968. Forty three years later, in 2001, this book was reprinted in paperback [52]. Twelve years on, in 2013, a second edition was printed [53].

However, advances in computer technology over the last decade or two have facilitated the widespread application of estimation theory. The widespread application was paralleled by a proliferation of textbooks written for a wide variety of readers, many of which did not have a strong background in statistical mathematics. The rigor of the mathematics as well as the notational systems of symbols and terminology vary considerably across the spectrum of the great number of textbooks and even more across the massive amount of web-based material available today. In an attempt to reduce confusion caused by the many reformulations of estimation theory, a specific formulation will be used going forward and that formulation, which is based on [30, 51], will be briefly explained.

The application of estimation theory involves the manipulation of random variables that are defined on a sample space $\mathcal{S}$. The manipulation takes into account that some of those random variables can be observed and some can not. The intermediate objective, if not the final objective, in virtually all estimation problems is to find the most compact density function for the random variable that represents a parameter of interest by using the information obtained by observing related random variables.

The probability density function can be used to find the probability that the outcome will map $x(\boldsymbol{\zeta})$ to a specified interval. For interval $x_1 \leq x(\zeta) \leq x_2$ that probability is given by

$$P(\{\zeta : x_1 \leq x(\zeta) \leq x_2\}) = \int_{x_1}^{x_2} f_x(x)\mathrm{d}x.$$

The joint probability density function can be used to find the probability that the outcome will map the target's coordinate, $(x(\zeta), y(\zeta))$ to a specified region. For region $\mathcal{R}$ that probability is

$$P(\{\zeta : (x(\zeta), y(\zeta)) \in \mathcal{R}\}) = \iint_{\mathcal{R}} f_{x,y}(x, y)\mathrm{d}x\mathrm{d}y.$$

A point estimator will be a constant whose value is obtained through mathematical manipulation of the probability density function. If the density function is described by a closed-form expression, then the constant will be a function of the parameters of the density function. For example, the estimator for a random variable with a Gaussian probability density function will be a function of its mean, μ, and variance, σ^2.

The MMSE point estimator is the constant x_{MMSE} that minimizes the probabilistic average of $(x(\boldsymbol{\zeta}) - x_{\mathrm{MMSE}})^2$. That is x_{MMSE} is the constant that minimizes the expectation

$$\mathbb{E}[(x(\boldsymbol{\zeta}) - x_{\mathrm{MMSE}})^2] = \int_{-\infty}^{\infty} (x - x_{\mathrm{MMSE}})^2 f_x(x)\mathrm{d}x. \tag{3.1}$$

This constant is given by

$$x_{\mathrm{MMSE}} = \mathbb{E}[x(\boldsymbol{\zeta})], \tag{3.2}$$

which is the mean value of $x(\boldsymbol{\zeta})$.

If another random variable, say $u(\boldsymbol{\zeta})$, is observed to have value $u(\zeta) = u^{\ddagger}$, then the information in the observation of $u(\boldsymbol{\zeta})$ usually rules out certain outcomes. If $u(\boldsymbol{\zeta})$ is in some way related to $x(\boldsymbol{\zeta})$ the information gained through its observation will change the density function of $x(\boldsymbol{\zeta})$. The modified density function for $x(\boldsymbol{\zeta})$ is referred to as the conditional density function of $x(\boldsymbol{\zeta})$ given $u(\zeta) = u$. This conditional density function is referred to as the *a posteriori* density function of $x(\boldsymbol{\zeta})$. It will be a function of u, which is the observed value of $u(\boldsymbol{\zeta})$. The conditional density function is symbolized $f_{x|u}(x \mid u)$. The first symbol in the subscript, which is the letter x, indicates the density function is for the random variable $x(\boldsymbol{\zeta})$ and the second indicates the random variable that was observed is $u(\boldsymbol{\zeta})$. The argument of the conditional density has two symbols separated by a vertical line. Both symbols represent real numbers. The first letter, which is x, is the true argument of the conditional density. The second symbol indicates the observed random variable is known to have value u.

In estimation scenarios where random variables are observed the *a posteriori* for a random variable takes into account that certain outcomes are no longer possible. Therefore, the *a posteriori* pdf provides the most up to date information about the random variable of interest. Therefore, the *a posteriori* density should be used to find the probability that the random variable is in a region. It is also used to find the MMSE estimate and the maximum *a posteriori* (MAP) point estimate, which is the point estimate given by the value of x that maximizes the *a posteriori* pdf.

Other estimators are used for reasons of computational efficiency. These estimators are usually sub-optimum with respect to the minimum mean squared error criterion. They could approximate the MMSE estimate or they could use some other cost criterion entirely, for example "minimize the maximum possible error".

---

‡Knowing that the outcome of the experiment is such that $u(\zeta) = u$ does not imply the outcome is known. This knowledge indicates the outcome is one of possibly many where $u(\zeta) = u$. The statement $u(\boldsymbol{\zeta}) = u$ has a slightly different meaning. It implies the experiment has not yet been performed, but once the experiment is performed the outcome ζ will be such that $u(\zeta) = u$.

Table 3.1: Identities for conditional probability functions.

No.	Identity
1	$f_{z,u}(z,u) = f_{z\|u}(z \mid u) f_u(u) = f_{u\|z}(u \mid z) f_z(z)$
2	$f_u(u) = \int_{-\infty}^{\infty} f_{z,u}(z,u)dz = \int_{-\infty}^{\infty} f_{u\|z}(u \mid z) f_z(z)dz$
3	$f_{z\|u}(z \mid u) = \frac{f_{u\|z}(u \mid z) f_z(z)}{f_u(u)}$; Bayes theorem
4	$f_{z\|u}(z \mid u) = \frac{f_{u\|z}(u \mid z) f_z(z)}{\int_{-\infty}^{\infty} f_{u\|z}(u \mid z) f_z(z)dz}$; more useful form of Bayes theorem
5	$f_{z,\, u\|q}(z,\ u \mid q) = f_{z\|u,\, q}(z \mid u,\ q) f_{u\|q}(u \mid q)$; chain rule or product rule

3.2 Developing the *A Posteriori* Probability Density Function

Since the *a posteriori* probability density function (pdf) is used to find the probability that the target is in a specified region as well as to find the MMSE estimate, the task at hand is to find the *a posteriori* pdf for target coordinates $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$. This is done using Bayes theorem, which is well known and fully explained in textbooks covering estimation theory [30, 51, 54, 55]. The theorem as well as other identities involving conditional density functions are summarized in Table 3.1.

The errors in the measurements of the gateways' locations can and will be approximated in a way that isolates their effect on the estimate of the target's location and, at the same time, greatly reduces the computation time needed to find the estimate. However, doing so comes at the expense of some accuracy, but the degradation will be tolerable and most of the time it will be insignificant. The quantity that will eventually be approximated is the range error, also referred to as range difference for Gateway ℓ, which is defined by:

$$\Delta r_\ell(\zeta) = \sqrt{\big(x_\ell(\zeta) - x(\zeta)\big)^2 + \big(y_\ell(\zeta) - y(\zeta)\big)^2} - \sqrt{\big(\widehat{x}_\ell(\zeta) - x(\zeta)\big)^2 + \big(\widehat{y}_\ell(\zeta) - y(\zeta)\big)^2}, \quad (3.3)$$

where $\widehat{x}_\ell(\zeta)$ and $\widehat{y}_\ell(\zeta)$ are the measured coordinates of Gateway ℓ for outcome ζ.

The objective of introducing $\Delta r_\ell(\zeta)$ is to isolate the effect of gateway location measurement error on the estimate of the target. To do that $\Delta r_\ell(\zeta)$ has to be transformed so that it is a random variable dependent only on $\Delta x_\ell(\zeta)$ and $\Delta y_\ell(\zeta)$.

Such an expression is obtained by manipulating $\sqrt{\big(x_\ell(\zeta)-x(\zeta)\big)^2+\big(y_\ell(\zeta)-y(\zeta)\big)^2}$ into a function of $\widehat{x}_\ell(\zeta)$ and $\widehat{y}_\ell(\zeta)$. First $x_\ell(\zeta)$ and $y_\ell(\zeta)$ are expressed as $\widehat{x}_\ell(\zeta)-\Delta x_\ell(\zeta)$ and $\widehat{y}_\ell(\zeta)-\Delta y_\ell(\zeta)$, respectively. Then the squared terms inside the square root are expanded to get

$$\begin{aligned}
&\sqrt{\big(x_\ell(\zeta)-x(\zeta)\big)^2+\big(y_\ell(\zeta)-y(\zeta)\big)^2}\\
&\quad=\sqrt{\big(\widehat{x}_\ell(\zeta)-\Delta x_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)+\Delta y_\ell(\zeta)-y(\zeta)\big)^2}\\
&\quad=\Big(\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2-2\Delta x_\ell(\zeta)\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)+\Delta x_\ell^2(\zeta)\\
&\qquad\quad+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2-2\Delta y_\ell(\zeta)\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)+\Delta y_\ell^2(\zeta)\Big)^{\frac{1}{2}}.
\end{aligned}\tag{3.4}$$

Assuming the error in the estimated location of Gateway ℓ is much smaller than the distance between the target and the gateway, i.e., assuming

$$\sqrt{\Delta x_\ell^2(\zeta)+\Delta y_\ell^2(\zeta)}\ll\sqrt{\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2},\tag{3.5}$$

then both $|\Delta x_\ell(\zeta)|$ and $|\Delta y_\ell(\zeta)|$ are much less than $\sqrt{\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2}$. Under this assumption, $\Delta x_\ell^2(\zeta)$ and $\Delta y_\ell^2(\zeta)$ in (3.4) are insignificant and can be removed. Removing the insignificant terms from (3.4) and re-arranging the argument of the square root produces the approximation

$$\begin{aligned}
&\sqrt{\big(x_\ell(\zeta)-x(\zeta)\big)^2+\big(y_\ell(\zeta)-y(\zeta)\big)^2}\simeq\\
&\qquad\sqrt{\Big(\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2\Big)\Big(1+2\epsilon_\ell(\zeta)\Big)},
\end{aligned}\tag{3.6}$$

where

$$\epsilon_\ell(\zeta)=-\frac{\Delta x_\ell(\zeta)\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)+\Delta y_\ell(\zeta)\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)}{\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2}.\tag{3.7}$$

Under the assumption for $\Delta x_\ell(\zeta)$ and $\Delta y_\ell(\zeta)$ expressed by (3.5), it can be shown that

$$\begin{aligned}
|\Delta x_\ell(\zeta)\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)|&\ll\ \big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{x}_\ell(\zeta)-y(\zeta)\big)^2\\
|\Delta y_\ell(\zeta)\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)|&\ll\ \big(\widehat{y}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2.
\end{aligned}\tag{3.8}$$

The inequalities of (3.8) ensure $\epsilon_\ell(\zeta)$ is much less than 1, which implies $\epsilon_\ell^2(\zeta)$ is much less than $2\epsilon_\ell(\zeta)$. The smallness of $\epsilon_\ell(\zeta)$ allows the term $\big(1+2\epsilon_\ell(\zeta)\big)$ to be replaced with $\big(1+2\epsilon_\ell(\zeta)+\epsilon_\ell^2(\zeta)\big)$ without adding significant error. Since $\big(1+2\epsilon_\ell(\zeta)+\epsilon^2(\zeta)\big) = \big(1+\epsilon_\ell(\zeta)\big)^2$, the modified (3.6) becomes

$$\sqrt{\big(x_\ell(\zeta)-x(\zeta)\big)^2+\big(y_\ell(\zeta)-y(\zeta)\big)^2} \simeq \big(1+\epsilon_\ell(\zeta)\big)\sqrt{\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2}. \tag{3.9}$$

Rearranging (3.9) produces the range difference approximation

$$\begin{aligned}\sqrt{\big(x_\ell(\zeta)-x(\zeta)\big)^2+\big(y_\ell(\zeta)-y(\zeta)\big)^2} - \sqrt{\big(\widehat{x}_\ell(\zeta)-x(\zeta)\big)^2+\big(\widehat{y}_\ell(\zeta)-y(\zeta)\big)^2} \\ \simeq \epsilon_\ell(\zeta)\sqrt{\big(x_\ell(\zeta)-x(\zeta)\big)^2+\big(y_\ell(\zeta)-y(\zeta)\big)^2}.\end{aligned} \tag{3.10}$$

Since the left side of (3.10) is $\Delta r_\ell(\zeta)$, substituting (3.7) for $\epsilon_\ell(\zeta)$ and changing (3.10), which is an equation with post experiment observations, to a pre-experiment equation has

$$\Delta r_\ell(\boldsymbol{\zeta}) \simeq -\frac{\Delta x_\ell(\boldsymbol{\zeta})\big(x_\ell(\boldsymbol{\zeta})-x(\boldsymbol{\zeta})\big)+\Delta y_\ell(\boldsymbol{\zeta})\big(y_\ell(\boldsymbol{\zeta})-y(\boldsymbol{\zeta})\big)}{\sqrt{\big(x_\ell(\boldsymbol{\zeta})-x(\boldsymbol{\zeta})\big)^2+\big(y_\ell(\boldsymbol{\zeta})-y(\boldsymbol{\zeta})\big)^2}}. \tag{3.11}$$

Since $\Delta x_\ell(\boldsymbol{\zeta})$ and $\Delta y_\ell(\boldsymbol{\zeta})$ are independent, zero-mean Gaussian random variables with the same variance, i.e., $\sigma_{x_\ell}^2 = \sigma_{y_\ell}^2 = \sigma_\ell^2$, it is easily shown that $\Delta r_\ell(\boldsymbol{\zeta})$ is Gaussian with zero-mean and variance $\sigma_{\Delta r_\ell}^2 = \sigma_\ell^2$. The proof of this is outlined below.

Proof of the variance of $\Delta r_\ell(\boldsymbol{\zeta})$: The proof falls out of the expression for the conditional expectation. Since the conditional mean of $\Delta r_\ell(\boldsymbol{\zeta})$ is 0, i.e., since $\mathbb{E}[\Delta r_\ell(\boldsymbol{\zeta}) \;\big|\; x(\boldsymbol{\zeta}) = x,\ y(\boldsymbol{\zeta}) = y,\ \widehat{x}_\ell(\boldsymbol{\zeta}) = \widehat{x}_\ell,\ \widehat{y}_\ell(\boldsymbol{\zeta}) = \widehat{y}_\ell] = 0$

$$\begin{aligned}\sigma_{\Delta r_\ell}^2 &= \mathbb{E}[(\Delta r_\ell(\boldsymbol{\zeta}))^2 \;\big|\; x(\boldsymbol{\zeta}) = x,\ y(\boldsymbol{\zeta}) = y, \widehat{x}_\ell(\boldsymbol{\zeta}) = \widehat{x}_\ell,\ \widehat{y}_\ell(\boldsymbol{\zeta}) = \widehat{y}_\ell] \\ &= \frac{\mathbb{E}[\Delta x_\ell^2(\boldsymbol{\zeta})]\big(\widehat{x}_\ell - x\big)^2 + \mathbb{E}[\Delta y_\ell^2(\boldsymbol{\zeta})]\big(\widehat{y}_\ell - y\big)^2}{\big(\widehat{x}_\ell - x\big)^2 + \big(\widehat{y}_\ell - y\big)^2} \\ &= \frac{\sigma^2 x_\ell\big(\widehat{x}_\ell - x\big)^2 + \sigma^2 y_\ell\big(\widehat{y}_\ell - y\big)^2}{\big(\widehat{x}_\ell - x\big)^2 + \big(\widehat{y}_\ell - y\big)^2}.\end{aligned} \tag{3.12}$$

If $\sigma^2_{x_\ell} = \sigma^2_{y_\ell}$, then the right side of (3.12) reduces to σ^2_ℓ. Therefore, σ^2_ℓ is the conditional variance of $\Delta r_\ell(\boldsymbol{\zeta})$. Clearly, the conditional variance of $\Delta r_\ell(\boldsymbol{\zeta})$ does not depend on the observations x, y, $\widehat{x}_\ell$ or $\widehat{y}_\ell$. This means, σ^2_ℓ is the variance of $\Delta r_\ell(\boldsymbol{\zeta})$. Therefore, $\Delta r_\ell(\boldsymbol{\zeta})$ is a zero-mean Gaussian RV with variance $\sigma^2_{\Delta r_\ell} = \sigma^2_\ell$ regardless of observations $x(\zeta)$, $y(\zeta)$, $\widehat{x}_\ell(\zeta)$ and $\widehat{y}_\ell(\zeta)$. Therefore, $\Delta r_\ell(\boldsymbol{\zeta})$ is a zero-mean Gaussian random variable with variance σ^2_ℓ if $\sigma^2_{x_\ell} = \sigma^2_{y_\ell}$. The difference is its variance will depend on the observations.

Should $\sigma^2_{y_\ell}$ not be equal to $\sigma^2_{x_\ell}$ then $\sigma^2_{\Delta r_\ell}$ will depend on the observations of $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$, $\widehat{x}_\ell(\boldsymbol{\zeta})$ and $\widehat{y}_\ell(\boldsymbol{\zeta})$. In this case the variance of Δr_ℓ can not be calculated, but it can easily be shown (see Appendix D on page 107) that is bounded by $\sigma^2_{\Delta r_\ell} \leq \max(\sigma^2_{x_\ell}, \sigma^2_{y_\ell})$.

It is important to point out that under conditions where $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$, $\widehat{x}_\ell(\boldsymbol{\zeta})$ and $\widehat{y}_\ell(\boldsymbol{\zeta})$ are known, i.e., observed, $\Delta r_\ell(\boldsymbol{\zeta})$ will be the sum of two independent zero-mean Gaussian random variables so it will be a zero-mean Gaussian random variable whether or not $\sigma^2_{x_\ell}$ is equal to $\sigma^2_{y_\ell}$.

The equation critical to the analysis going forward is the equation that relates $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ to the range from Gateway ℓ to the target. This equation is obtained by updating (1.4) to use the nomenclature developed in the previous chapter. The updated (1.4) becomes

$$\widehat{\mathsf{ToF}}_\ell(\boldsymbol{\zeta}) = \frac{\sqrt{(x_\ell(\boldsymbol{\zeta}) - x(\boldsymbol{\zeta}))^2 + (y_\ell(\boldsymbol{\zeta}) - y(\boldsymbol{\zeta}))^2}}{c}. \tag{3.13}$$

Substituting the above for $\mathsf{ToF}_\ell(\boldsymbol{\zeta})$ in (2.3) produces

$$\begin{aligned}
\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta}) &= \mathsf{ToF}_\ell(\boldsymbol{\zeta}) + t(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta}) \\
&= \frac{\sqrt{(x_\ell(\boldsymbol{\zeta}) - x(\boldsymbol{\zeta}))^2 + (y_\ell(\boldsymbol{\zeta}) - y(\boldsymbol{\zeta}))^2}}{c} + t(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta}) \\
&= \frac{\sqrt{(\widehat{x}_\ell(\boldsymbol{\zeta}) - x(\boldsymbol{\zeta}))^2 + (\widehat{y}_\ell(\boldsymbol{\zeta}) - y(\boldsymbol{\zeta}))^2} + \Delta r_\ell(\boldsymbol{\zeta})}{c} + t(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta}) \\
&= \frac{\sqrt{(\widehat{x}_\ell(\boldsymbol{\zeta}) - x(\boldsymbol{\zeta}))^2 + (\widehat{y}_\ell(\boldsymbol{\zeta}) - y(\boldsymbol{\zeta}))^2}}{c} + \Delta\mathsf{ToF}_\ell(\boldsymbol{\zeta}) + t(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta}),
\end{aligned} \tag{3.14}$$

where $\Delta\mathsf{ToF}_\ell(\boldsymbol{\zeta}) = \dfrac{\Delta r_\ell(\boldsymbol{\zeta})}{c}$ represents the difference in the time-of-flight from the target

to the true and measured locations of Gateway ℓ, respectively, and $\Delta T_\ell(\boldsymbol{\zeta})$ represents the error in the time-of-arrival measurements. Both random variables are zero-mean Gaussian under conditions where $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$, $\widehat{x}_\ell(\boldsymbol{\zeta})$ and $\widehat{y}_\ell(\boldsymbol{\zeta})$ have been observed but the former has variances $\sigma^2_{x_\ell}$ and $\sigma^2_{y_\ell}$, which are σ^2_ℓ/c^2 if $\sigma_{x_\ell} = \sigma_{y_\ell}$, and the latter has variance $\sigma^2_{t_\ell}$. Each of $\Delta T_\ell(\boldsymbol{\zeta})$ is independent of all other fundamental RVs. Also, each of $\Delta\mathsf{ToF}_\ell(\boldsymbol{\zeta})$ is independent of all fundamental RVs except $\Delta x_\ell(\boldsymbol{\zeta})$ and $\Delta y_\ell(\boldsymbol{\zeta})$.

The problem at hand has three parameters that are neither known nor measured, but are related. These three parameters are modeled as RVs $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$, and $t(\boldsymbol{\zeta})$. The conditional joint densities of interest is that for $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$, but in this case the conditional joint density for all three RVs must be obtained and then integrated over t, where t is the variable in the argument of the density function for RV $t(\boldsymbol{\zeta})$, to produce the conditional joint density of $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$. That is to say, the desired density function is

$$f_{x,y\left|\widehat{\mathsf{ToA}}_\ell,\, \widehat{x}_\ell,\, \widehat{y}_\ell,\, \ell=1,2,\ldots,L\right.}(x, y \mid \widehat{\mathsf{ToA}}_\ell, \widehat{x}_\ell, \widehat{y}_\ell,\ \ell = 1, 2, \ldots, L), \tag{3.15}$$

which is obtained by the integration

$$\int_{-\infty}^{\infty} f_{x,y,t\left|\widehat{\mathsf{ToA}}_\ell,\, \widehat{x}_\ell,\, \widehat{y}_\ell,\, \ell=1,2,\ldots,L\right.}(x, y, t \mid \widehat{\mathsf{ToA}}_\ell, \widehat{x}_\ell, \widehat{y}_\ell,\ \ell = 1, 2, \ldots, L)dt. \tag{3.16}$$

The subscript and argument lists for the conditional density functions of (3.15) and (3.16) are quite long and cumbersome. Unfortunately, both the subscript and the argument contain essential information so neither can be omitted. In order to shorten the notation, the vectors defined in Table 3.2 are introduced to the nomenclature.

To help distinguish vectors from scalars, the symbols for vectors will be in bold font. The elements of a vector may be random variables or observations of random variables so the vector may be a function of $\boldsymbol{\zeta}$ or ζ. For example, the elements of vector $\mathbf{ToA}(\boldsymbol{\zeta})$ defined in Table 3.2 contains the RVs $\mathsf{ToA}_1(\boldsymbol{\zeta})$, $\mathsf{ToA}_2(\boldsymbol{\zeta})$, ..., $\mathsf{ToA}_L(\boldsymbol{\zeta})$. Similarly, $\mathbf{ToA}(\zeta)$ is a vector of observations with elements $\mathsf{ToA}_1(\zeta)$, $\mathsf{ToA}_2(\zeta)$, ..., $\mathsf{ToA}_L(\zeta)$. $\mathbf{ToA}(\zeta)$ is said to be an observation or instance of $\mathbf{ToA}(\boldsymbol{\zeta})$.

Table 3.2: Definition of relevant vectors.

Vector	Description	Definition
$\boldsymbol{g}_x(\boldsymbol{\zeta})$	Vector of RVs representing the true x coordinates for the L gateways	$\boldsymbol{g}_x(\boldsymbol{\zeta}) = \big[x_1(\boldsymbol{\zeta}),\ x_2(\boldsymbol{\zeta}),\ \ldots,\ x_L(\boldsymbol{\zeta})\big]^\top$
$\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})$	Vector of RVs representing the measured x coordinates for the L gateways	$\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \big[\widehat{x}_1(\boldsymbol{\zeta}),\ \widehat{x}_2(\boldsymbol{\zeta}),\ \ldots,\ \widehat{x}_L(\boldsymbol{\zeta})\big]^\top$
$\boldsymbol{g}_y(\boldsymbol{\zeta})$	Vector of RVs representing the true y coordinates for the L gateways	$\boldsymbol{g}_y(\boldsymbol{\zeta}) = \big[y_1(\boldsymbol{\zeta}),\ y_2(\boldsymbol{\zeta}),\ \ldots,\ y_L(\boldsymbol{\zeta})\big]^\top$
$\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})$	Vector of RVs representing the measured y coordinates for the L gateways	$\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \big[\widehat{y}_1(\boldsymbol{\zeta}),\ \widehat{y}_2(\boldsymbol{\zeta}),\ \ldots,\ \widehat{y}_L(\boldsymbol{\zeta})\big]^\top$
$\mathbf{ToA}(\boldsymbol{\zeta})$	Vector of RVs representing the times of arrival at the L gateways in the absence of all measurement errors	$\mathbf{ToA}(\boldsymbol{\zeta}) = [\mathsf{ToA}_1(\boldsymbol{\zeta}), \ldots, \mathsf{ToA}_L(\boldsymbol{\zeta})]^\top$
$\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$	Vector of RVs representing the measured times of arrival at the L gateways	$\widehat{\mathbf{ToA}}(\boldsymbol{\zeta}) = [\widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}), \ldots, \widehat{\mathsf{ToA}}_L(\boldsymbol{\zeta})]^\top$
$\boldsymbol{\tau}$	Vector of RVs representing the measured values of the times of arrival at the L gateways. This vector is an observation of $\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$, i.e., $\boldsymbol{\tau} = \widehat{\mathbf{ToA}}(\zeta)$	$\boldsymbol{\tau} = [\tau_1, \ldots, \tau_L]^\top = [\widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}), \ldots, \widehat{\mathsf{ToA}}_L(\boldsymbol{\zeta})]^\top$

The vectors defined in Table 3.2 allow the two conditional density functions in (3.15) and (3.16) to be written as

$$f_{x,y\left|\widehat{\mathsf{ToA}}_\ell,\, \widehat{x}_\ell,\, \widehat{y}_\ell,\, \ell=1,2,\ldots,L\right.}(x, y \mid \tau_\ell, \widehat{x}_\ell, \widehat{y}_\ell,\ \ell = 1, 2, \ldots, L) = f_{x,y\left|\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.}(x, y \mid \boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y) \quad (3.17)$$

and

$$f_{x,y,t\left|\widehat{\mathsf{ToA}}_\ell,\widehat{x}_\ell,\widehat{y}_\ell,\, \ell=1,2,\ldots,L\right.}(x, y, t \mid \widehat{\mathsf{ToA}}_\ell, \widehat{x}_\ell, \widehat{y}_\ell,\ \ell = 1, 2, \ldots, L) = f_{x,y,t\left|\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.}(x, y, t \mid \boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y). \quad (3.18)$$

Note that the vector $\boldsymbol{\tau}$, which appears in the arguments of the conditional densities on the right sides of (3.17) and (3.18), is implicitly defined to be the observation of $\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$. The vector $\boldsymbol{\tau}$ is used in place of $\widehat{\mathbf{ToA}}(\zeta)$ to make the notation even more compact.

Bayes theorem allows the latter joint density to be expressed as

$$f_{x,y,t\left|\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.}(x,y,t \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) = \frac{f_{\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\left|\right._{x,y,t}(\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y \mid x,y,t) f_{x,y,t}(x,y,t)}{f_{\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y}(\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)}. \tag{3.19}$$

Random variables $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$, while not necessarily independent of each other, are independent of $t(\boldsymbol{\zeta})$. This fact allows the joint density $f_{x,y,t}(x,y,t)$ to be expressed as

$$f_{x,y,t}(x,y,t) = f_{x,y}(x,y) f_t(t).$$

Making this substitution into (3.19) and then applying the theorem of total probability to the denominator on the right side of (3.19), i.e., the denominator is the triple integration from -∞ to ∞ of the numerator with respect to x, y and t produces

$$f_{x,y,t\left|\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.}(x,y,t\left|\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.) = \frac{f_{\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\left|\right._{x,y,t}(\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\left|x,y,t\right.) f_{x,y}(x,y) f_t(t)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\widehat{\mathbf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\left|\right._{x,y,t}(\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\left|x,y,t\right.) f_{x,y}(x,y) f_t(t) \mathrm{d}x\mathrm{d}y\mathrm{d}t}. \tag{3.20}$$

It can be argued, very convincingly, that $f_t(t)$ can be moved outside the triple integral in the denominator and cancel the $f_t(t)$ in the numerator. On the surface this seems to defy the basic rule of integration which says a function of t can not be moved outside an integral with respect to t. However, in this case, it is argued that $f_t(t)$ is constant over the interval of t where the integrand is non-zero.

The mathematical logic for moving $f_t(t)$ outside the triple integral is listed below

1. The density function $f_t(t)$ is given by

$$f_t(t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{t^2}{2\sigma^2}\right),$$

where σ is very large and approaches ∞. This implies $f_t(t) = \frac{1}{\sqrt{2\pi}\sigma}$ for any t with finite

value.

2. The physics dictates that signal strength limitation restricts the distance between a gateway and the target to be at most 100 km. This means the time-of-flight, which is $\mathsf{ToA}_\ell(\zeta) - \mathsf{t}(\zeta)$, will be less than $100\text{km}/c$.

3. Therefore
$$\mathsf{ToA}_\ell(\zeta) - 10^5 m/c < t(\zeta) < \mathsf{ToA}_\ell(\zeta),\ \ell = 1, 2, \ldots, L.$$

4. The time origin will certainly be defined so that $\mathsf{ToA}_\ell(\zeta)$ is finite, therefore the value of $t(\zeta)$ is also finite.

5. Since $f_t(t)$ is a zero mean Gaussian pdf with a variance that tends to ∞ the exponent will be nearly zero and $f_t(t) \to \frac{1}{\sqrt{2\pi}\sigma}$ for a finite value of $t(\zeta)$.

6. Therefore $f_t(t)$ is constant over the interval of t where the integrand is possibly non-zero and can be moved outside the integral with respect to t.

After moving $f_t(t)$ outside the triple integral and canceling the same term in the numerator, Bayes equation becomes

$$f_{x,\,y,\,t\left|\widehat{\mathsf{ToA}},\,\widehat{\boldsymbol{g}}_x,\,\widehat{\boldsymbol{g}}_y\right.}(x, y, t \mid \boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y) = \frac{f_{x,y}(x,y) f_{\widehat{\mathsf{ToA}},\,\widehat{\boldsymbol{g}}_x,\,\widehat{\boldsymbol{g}}_y\left|x,\,y,\,t\right.}(\boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y \mid x, y, t)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}(x,y) f_{\widehat{\mathsf{ToA}},\,\widehat{\boldsymbol{g}}_x,\,\widehat{\boldsymbol{g}}_y\left|x,y,t\right.}(\boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y \mid x, y, t)\mathrm{d}x\mathrm{d}y\mathrm{d}t}. \tag{3.21}$$

The right side of (3.21) can be further simplified using the fact that $\widehat{x}_\ell(\boldsymbol{\zeta})$, $\widehat{y}_\ell(\boldsymbol{\zeta})$ and $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ are independent of $\widehat{x}_k(\boldsymbol{\zeta})$, $\widehat{y}_k(\boldsymbol{\zeta})$ and $\widehat{\mathsf{ToA}}_k(\boldsymbol{\zeta})$ for $\ell \neq k$. This independence relationship allows the conditional joint density to be factored into the conditional joint densities

$f_{\widehat{\mathsf{ToA}}_\ell,\,\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\tau_\ell,\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t)$, $\ell=1,2,\ldots,L$, with the result

$$f_{x,\,y,\,t\big|\widehat{\mathbf{ToA}},\,\widehat{\boldsymbol{g}}_x,\,\widehat{\boldsymbol{g}}_y}(x,y,t \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) = \frac{f_{x,y}(x,y)\prod_{\ell=1}^{L} f_{\widehat{\mathsf{ToA}}_\ell,\,\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\tau_\ell,\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}(x,y)\prod_{\ell=1}^{L} f_{\widehat{\mathsf{ToA}}_\ell,\,\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\tau_\ell,\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t)dxdydt}. \tag{3.22}$$

The reformation of (3.19) is not yet complete. It is transformed to a useful form in 5 steps.

Step 1: The conditional joint density $f_{\widehat{\mathsf{ToA}}_\ell,\,\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\tau_\ell,\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t)$ is transformed using the chain rule, which is given in the 5$^{\text{th}}$ entry of Table 3.1, to get

$$f_{\widehat{\mathsf{ToA}}_\ell,\,\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\tau_\ell,\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t) = f_{\widehat{\mathsf{ToA}}_\ell\big|\widehat{x}_\ell,\,\widehat{y}_\ell,\,x,\,y,\,t}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) f_{\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t). \tag{3.23}$$

Step 2: Since $\widehat{x}_\ell(\boldsymbol{\zeta})$ and $\widehat{y}_\ell(\boldsymbol{\zeta})$ are independent of $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$ and $t(\boldsymbol{\zeta})$, $f_{\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t)$ is equal to $f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)$ and therefore the latter can replace the former in (3.22).

Step 3: Substituting (3.23) into (3.22) and then replacing $f_{\widehat{x}_\ell,\,\widehat{y}_\ell\big|x,\,y,\,t}(\widehat{x}_\ell,\widehat{y}_\ell \mid x,y,t)$ with $f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)$ transforms (3.22), which means it also transforms (3.21), into

$$f_{x,\,y,\,t\big|\widehat{\mathbf{ToA}},\,\widehat{\boldsymbol{g}}_x,\,\widehat{\boldsymbol{g}}_y}(x,y,t \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) = \frac{f_{x,y}(x,y)\prod_{\ell=1}^{L} f_{\widehat{\mathsf{ToA}}_\ell\big|\widehat{x}_\ell,\,\widehat{y}_\ell,\,x,\,y,\,t}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}(x,y)\prod_{\ell=1}^{L} f_{\widehat{\mathsf{ToA}}_\ell\big|\widehat{x}_\ell,\,\widehat{y}_\ell,\,x,\,y,\,t}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)\mathrm{d}x\mathrm{d}y\mathrm{d}t}. \tag{3.24}$$

Step 4: The quantity $\prod_{\ell=1}^{L} f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)$ can be taken outside the triple integral in the denominator where it cancels the same quantity in the numerator, since $f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)$ does not depend on x, y or t. After removing $\prod_{\ell=1}^{L} f_{\widehat{x}_\ell,\,\widehat{y}_\ell}(\widehat{x}_\ell,\widehat{y}_\ell)$ from both the numerator and

denominator, (3.22), and therefore (3.21) as well, becomes

$$f_{x,y,t\left|\widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.}(x,y,t\big|\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)0$$

$$= \frac{f_{x,y}(x,y)\prod\limits_{\ell=1}^{L} f_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell\big|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t)}{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} f_{x,y}(x,y)\prod\limits_{\ell=1}^{L} f_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell\big|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t)dxdydt}. \tag{3.25}$$

Step 5: Finally the conditional density $f_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\,\widehat{y}_\ell,\,x,\,y,\,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t)$ is evaluated. This is done using first principles to evaluate the associated cumulative distribution function, which is $F_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\,\widehat{y}_\ell,\,x,\,y,\,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t)$. Then $f_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\,\widehat{y}_\ell,\,x,\,y,\,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t)$ is obtained from the distribution function. It will be shown that this conditional distribution function is Gaussian, which means once the mean and variance are obtained the conditional joint density is also obtained.

The conditional distribution function, by definition, is given by

$$F_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) =$$
$$P\big(\{\zeta : \widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta}) \le \tau_\ell \mid \widehat{x}_\ell(\boldsymbol{\zeta}) = \widehat{x}_\ell,\ \widehat{y}_\ell(\boldsymbol{\zeta}) = \widehat{y}_\ell,\ x(\boldsymbol{\zeta}) = x,\ y(\boldsymbol{\zeta}) = y,\ t(\boldsymbol{\zeta}) = t\}\big).$$

The primary RV is $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ so by definition the set of $\zeta_i's$ in the argument of the probability function can also be defined in terms of the RVs on the right side of (3.14) by substituting the right side of (3.14) for $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$. Making this substitution and then substituting the given conditions into the result has

$$F_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) =$$
$$P\big(\{\zeta : t + \frac{\sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}}{c} + \Delta\mathsf{ToF}_\ell(\boldsymbol{\zeta}) + \Delta T_\ell(\boldsymbol{\zeta}) \le \tau_\ell\}\big). \tag{3.26}$$

Rearranging terms in the inequality that defines the set of $\zeta_i's$ in (3.26) produces

$$\begin{aligned} &F_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) = \\ &\quad P\big(\{\zeta : \Delta T_\ell(\boldsymbol{\zeta}) + \Delta\mathsf{ToF}_\ell(\boldsymbol{\zeta}) \le \tau_\ell - \frac{\sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}}{c} - t\}\big), \end{aligned} \tag{3.27}$$

where, to make the equation more compact, $\widehat{x}_\ell$, $\widehat{y}_\ell$, t and τ_ℓ are used in place of $\widehat{x}_\ell(\zeta)$, $\widehat{y}_\ell(\zeta)$, $t(\zeta)$ and $\tau_\ell(\zeta)$, respectively.

Since the observed values of $x(\boldsymbol{\zeta})$, $y(\boldsymbol{\zeta})$, $\widehat{x}_\ell(\boldsymbol{\zeta})$ and $\widehat{y}_\ell(\boldsymbol{\zeta})$ are used in (3.26), $\Delta\mathsf{ToF}(\boldsymbol{\zeta})$, which is located on the left side of the set defining inequality, will be a zero-mean Gaussian random variable. Therefore, the left side of the set defining inequality in (3.27) is the sum of 2 independent zero-mean Gaussian random variables. Let the sum of those RVs be denoted $\mathsf{sum}_\ell(\boldsymbol{\zeta})$, then $\mathsf{sum}_\ell(\boldsymbol{\zeta}) = \Delta T_\ell(\boldsymbol{\zeta}) + \Delta\mathsf{ToF}_\ell(\boldsymbol{\zeta})$ and $\mathsf{sum}_\ell(\boldsymbol{\zeta})$ is a zero-mean Gaussian random variable with variance

$$\sigma^2_{\mathsf{sum}_\ell} = \sigma^2_{t_\ell} + \sigma^2_{\Delta r_\ell}/c^2,$$

where $\sigma^2_{\Delta r_\ell} = \sigma^2_\ell$ if $\sigma^2_{x_\ell} = \sigma^2_{y_\ell} = \sigma^2_\ell$. This fact allows (3.27) to be expressed as

$$\begin{aligned} &F_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) = \\ &\qquad P\big(\{\zeta : \mathsf{sum}_\ell(\boldsymbol{\zeta}) \le \tau_\ell - \frac{\sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}}{c} - t\}\big). \end{aligned}$$

By definition the probability on the right side is $F_{\mathsf{sum}_\ell}(\mathsf{sum}_\ell)\big|_{\mathsf{sum}_\ell = \tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t}$. Therefore,

$$F_{\widehat{\mathsf{ToA}}_\ell\left|\widehat{x}_\ell,\widehat{y}_\ell,x,y,t\right.}(\tau_\ell \mid \widehat{x}_\ell,\widehat{y}_\ell,x,y,t) = F_{\mathsf{sum}_\ell}(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t), \tag{3.28}$$

where, as just explained, $F_{\mathsf{sum}_\ell}(\mathsf{sum}_\ell)$ is a zero-mean Gaussian distribution with variance $\sigma^2_{\mathsf{sum}_\ell}$.

Having established the conditional distribution function for $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$, it is clear that

the conditional density function for $\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ is given by

$$\begin{aligned} & f_{\widehat{\mathsf{ToA}}_\ell \left| \widehat{x},\widehat{y},x,y,t \right.}(\tau_\ell \mid \widehat{x},\widehat{y},x,y,t) \\ & \quad = f_{\mathsf{sum}_\ell}\Big(\tau_\ell - \frac{\sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}}{c} - t\Big) \\ & \quad = \frac{1}{\sqrt{2\pi\sigma^2_{\mathsf{sum}_\ell}}} \exp\Big(\frac{(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t)^2}{2\sigma^2_{\mathsf{sum}_\ell}}\Big). \end{aligned} \tag{3.29}$$

Substituting (3.29) into (3.25) produces

$$\begin{aligned} & f_{x,y,t \left| \widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y \right.}(x,y,t \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) \\ & = \frac{f_{x,y}(x,y)\prod_{\ell=1}^L (2\pi\sigma^2_{\mathsf{sum}_\ell})^{-\frac{1}{2}} \exp\Big(-\frac{(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t)^2}{2\sigma^2_{\mathsf{sum}_\ell}}\Big)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}(x,y)\prod_{\ell=1}^L (2\pi\sigma^2_{\mathsf{sum}_\ell})^{-\frac{1}{2}} \exp\Big(-\frac{(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t)^2}{2\sigma^2_{\mathsf{sum}_\ell}}\Big)\mathrm{d}x\mathrm{d}y\mathrm{d}t} \\ & = \frac{f_{x,y}(x,y)\prod_{\ell=1}^L \exp\Big(-\frac{(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t)^2}{2\sigma^2_{\mathsf{sum}_\ell}}\Big)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{x,y}(x,y)\prod_{\ell=1}^L \exp\Big(-\frac{(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c - t)^2}{2\sigma^2_{\mathsf{sum}_\ell}}\Big)\mathrm{d}x\mathrm{d}y\mathrm{d}t}. \end{aligned} \tag{3.30}$$

Equation (3.30) is a useful expression for $f_{x,y,t \left| \widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y \right.}(x,y,t \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)$ but the objective is to find a useful expression for $f_{x,y \left| \widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y \right.}(x,y \big| \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)$. The conditional joint density for $f_{x,y \left| \widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y \right.}(x,y \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)$ is obtained by integrating $f_{x,y,t \left| \widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y \right.}(x,y,t \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)$ with respect to t from $-\infty$ to ∞. Careful examination of (3.30) shows that the denominator on the right side is already integrated with respect to t, which means it is not a function of t. However, the numerator is a function of t and needs to be integrated with respect to t. It also shows that the numerator is $f_{x,y}(x,y)$ multiplied by an exponential function of t with exponent

$$-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\big(t - (\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c)\big)^2}{(\sigma^2_{\mathsf{sum}_\ell})} = -\frac{1}{2}\sum_{\ell=1}^{L}\frac{\big(t - \mu_\ell\big)^2}{(\sigma^2_{\mathsf{sum}_\ell})}, \tag{3.31}$$

where $\mu_\ell = \tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c$.

Clearly, the exponent is the sum of quadratics in t, which will be a quadratic in t. This means the integration with respect to t can be performed symbolically. The method for this integration has two steps: the first step, which is easily done, but very tedious, is to manipulate the exponent into the form $-\frac{(t-\mu)^2}{2\sigma^2} - \frac{1}{2}\gamma$, where μ, σ^2, and γ are not functions of t. Then the integration with respect to t is quite simple since it is well known that $\int_{-\infty}^{\infty} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt = \sqrt{2\pi\sigma^2}$. Integrating both the numerator and denominator in the right side of (3.30) with respect to t has

$$f_{x,y}\Big|_{\widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y}(x,y \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) = \frac{\sqrt{2\pi\sigma^2}e^{-\frac{\gamma}{2}}f_{x,y}(x,y)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\sqrt{2\pi\sigma^2}e^{-\frac{\gamma}{2}}f_{x,y}(x,y)\mathrm{d}x\mathrm{d}y}. \tag{3.32}$$

Since σ^2 does not depend on x or y, $\sqrt{2\pi\sigma^2}$ can be taken outside the double integral in the denominator. In doing so it cancels with the same term in the numerator leaving

$$f_{x,y}\Big|_{\widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y}(x,y \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) = \frac{e^{-\frac{\gamma}{2}}f_{x,y}(x,y)}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-\frac{\gamma}{2}}f_{x,y}(x,y)\mathrm{d}x\mathrm{d}y}. \tag{3.33}$$

The tedious manipulation of (3.31) to get the form $-(t-\mu)^2/(2\sigma^2) - 1/\gamma$ has been relegated to Appendix A on page 102. It shows that μ, $\frac{1}{\sigma^2}$, and γ are given by

$$\mu = \frac{\sum_{\ell=1}^{L}\frac{\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c}{\sigma^2_{\mathsf{sum}_\ell}}}{\sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}}, \tag{3.34}$$

$$\frac{1}{\sigma^2} = \sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}, \tag{3.35}$$

$$\begin{aligned}\gamma &= \sum_{\ell=1}^{L}\frac{\left(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c\right)^2}{\sigma^2_{\mathsf{sum}_\ell}} \\ &- \left(\frac{1}{\sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}}\right)\left(\sum_{\ell=1}^{L}\frac{\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}/c}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2,\end{aligned} \tag{3.36}$$

$$\sigma^2_{\mathsf{sum}_\ell} = \sigma^2_{t_\ell} + \sigma^2_{\Delta r_\ell}/c^2 = \sigma^2_{t_\ell} + \frac{\sigma^2_{x_\ell}\left(\widehat{x}_\ell - x\right)^2 + \sigma^2_{y_\ell}\left(\widehat{y}_\ell - y\right)^2}{\left(\widehat{x}_\ell - x\right)^2 + \left(\widehat{y}_\ell - y\right)^2}. \tag{3.37}$$

Note that the denominator on the right side of (3.33) is a constant that makes the volume under $\exp(-\gamma/2)\ f_{x,y}(x,y)$ equal to 1. The double integral that produces the constant will likely have to be performed numerically.

3.3 Defining the *a Priori* PDF

The initial *a priori* pdf, $f_{x,y}(x,y)$, is approximated using physics, the geography and knowing the region where the sensors were deployed. While physics may allow the exact calculation of $f_{x,y}(x,y)$, practically the mathematics are usually too complicated to find $f_{x,y}(x,y)$ exactly.

For example, suppose the class A target is known to be inoperable in water. Then physics dictates that $f_{x,y}(x,y)$ should be zero for x,y coordinates where water is covering the surface of the earth. The problem is the shorelines of oceans, lakes and ponds change hourly depending on the tide, wind and rain, making it practically impossible to zero out the correct regions of $f_{x,y}(x,y)$. Practically, it would make sense to only consider large bodies of water and then either use their "average" shorelines or include a taper in $f_{x,y}(x,y)$ to account for the uncertainty in the shoreline.

In many applications no data are available to support the physics and mathematics used to find *a priori* pdf $f_{x,y}(x,y)$. In such cases $f_{x,y}(x,y)$ is assumed to be widely uniform. For example, $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ are assumed to be independent zero-mean Gaussian random variables with variances that tend to ∞. Under this assumption $f_{x,y}(x,y)$ can be taken outside the double integral in the denominator on the right side of (3.33) for the same reason that $f_t(t)$ was taken outside the triple integral. Once taken outside the integral it cancels $f_{x,y}(x,y)$ in the numerator.

There may be occasions when the joint *a posteriori* pdf for $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ has been found from other measurements that were made previous to the time of arrival measurements. Should that be the case, the *a posteriori* found with those measurements becomes the *a priori* pdf in (3.32). In fact the *a priori* pdf is found using (3.32) can be used as the *a priori* pdf the next time (3.32) is used. For example, suppose it is known that the target is fixed.

On the first transmission the *a posteriori* pdf is found using (3.33) with the $\widehat{\mathbf{ToA}}$ and an approximate *a priori* pdf that was found based on physics. Then, on a second transmission, the *a posteriori* pdf is updated using the *a posteriori* pdf found on the first transmission for the *a priori* pdf.

3.4 The MMSE and MAP Estimators for $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$

It must be emphasized that the primary objective of this work is to find the probability that the target is in a specified credible region[‡]. However, since the literature only discusses points estimator for time of arrival based localization applications, the discussion on point estimators has been disproportionate. The MMSE estimate for $x(\boldsymbol{\zeta})$ given the observations of $\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$, $\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})$ and $\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})$ is

$$x_{\text{MMSE}}(\zeta) = \mathbb{E}[x(\boldsymbol{\zeta}) \mid \widehat{\mathbf{ToA}}(\boldsymbol{\zeta}) = \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y]. \tag{3.38}$$

Obviously, the estimate $x_{\text{MMSE}}(\zeta)$ depends on the outcome of the experiment through the observations of $\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$, $\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})$ and $\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})$, i.e, through $\boldsymbol{\tau}$, $\widehat{\boldsymbol{g}}_x$ and $\widehat{\boldsymbol{g}}_x$. This means $x_{\text{MMSE}}(\zeta)$ can be viewed as an observation of random variable $x_{\text{MMSE}}(\boldsymbol{\zeta})$. Random variable $x_{\text{MMSE}}(\boldsymbol{\zeta})$ is referred to as the MMSE estimator. The estimator is obtained by substituting $\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$ for $\boldsymbol{\tau}$, $\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})$ for $\widehat{\boldsymbol{g}}_x$, $\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})$ for $\widehat{\boldsymbol{g}}_y$ i.e.,

$$x_{\text{MMSE}}(\boldsymbol{\zeta}) = \mathbb{E}[x(\boldsymbol{\zeta})\big|\widehat{\mathbf{ToA}}(\boldsymbol{\zeta}) = \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y] \Big|_{\boldsymbol{\tau}=\widehat{\mathbf{ToA}}(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_x=\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_y=\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})}. \tag{3.39}$$

The expectation in both (3.38) and (3.39) is taken using the conditional density

$$f_{x\left|\widehat{\mathbf{ToA}},\,\widehat{\boldsymbol{g}}_x,\,\widehat{\boldsymbol{g}}_y\right.}(x \mid \boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y).$$

[‡]In Bayesian analysis such a region is called a credible region as opposed to a confidence region, which is the term used for a region in statistical analysis.

Substituting $\widehat{\mathbf{ToA}}(\boldsymbol{\zeta})$ for $\boldsymbol{\tau}$, $\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})$ for $\widehat{\boldsymbol{g}}_x$, $\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})$ for $\widehat{\boldsymbol{g}}_y$ produces

$$x_{\text{MMSE}}(\boldsymbol{\zeta}) = \left. \frac{\int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} e^{-\frac{\gamma}{2}} dy dx}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{\gamma}{2}} dy dx} \right|_{\boldsymbol{\tau}=\widehat{\mathbf{ToA}}(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_x=\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_y=\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})} . \tag{3.40}$$

Similarly, the MMSE estimator for $y(\boldsymbol{\zeta})$ is given by

$$\begin{aligned} y_{\text{MMSE}}(\boldsymbol{\zeta}) &= \mathbb{E}[y(\boldsymbol{\zeta}) \mid \widehat{\mathbf{ToA}}(\boldsymbol{\zeta}) = \boldsymbol{\tau},\, \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\, \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y] \mid_{\boldsymbol{\tau}=\widehat{\mathbf{ToA}}(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_x=\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_y=\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})} \\ &= \left. \frac{\int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} e^{-\frac{\gamma}{2}} dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{\gamma}{2}} dy dx} \right|_{\boldsymbol{\tau}=\widehat{\mathbf{ToA}}(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_x=\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}),\, \widehat{\boldsymbol{g}}_y=\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})} . \end{aligned} \tag{3.41}$$

Unfortunately, the integrals on the right sides of (3.40) and (3.41) are extremely difficult if not impossible to evaluate symbolically and most likely will have to be done numerically. Even with efficient numerical integration, the computation time required to evaluate (3.40) and (3.41) may be significant.

The joint maximum *a posteriori* (MAP) estimates are easier to obtain. They are the values of x and y that jointly maximizes (3.33). Since the denominator on the right side of (3.33) does not depend on either x or y, it plays no role in the maximization and can be removed. It is pointed out that the numerator by itself is not a density as the volume under its surface is not 1. The numerator is referred to as the maximum *a posteriori* function, i.e., the MAP function. In the special case where the *a priori* joint density, which is $f_{x,y}(x, y)$ in (3.33) is constant over the region of x and y where $e^{-\frac{\gamma}{2}}$ is non-zero, the MAP function is also the maximum likelihood (ML) function and the MAP estimator is also the ML estimator.

Since an exponential function is monotonic increasing with its argument, the values of x and y that maximize $e^{-\frac{\gamma}{2}}$ also maximize $-\gamma$. Therefore, the ML estimate is the values of x and y that maximize $-\gamma$ or minimize γ. This is also the case for the MAP estimate if

$f_{x,y}(x, y)$ is widely flat.

Since γ is a complicated function of x and y, the ML estimator for $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ is found numerically using a hill climbing or valley descending algorithm. Such algorithms are much more computationally efficient than numerical integration algorithms.

Valley descending algorithms perform much better if the gradient is available in closed form. For this reason a closed form expression for the gradients associated with two likelihood functions, both denoted $f_{\mathsf{ToA}|\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y,x,y}(\boldsymbol{\tau} \mid \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y, x, y)$, one for t known and the other for t unknown are developed in Appendix B on page 104. The results are given in equations (B.1) and (B.2), respectively.

The complex computation of the integrals on the right sides of (3.40) and (3.41) can be avoided at the expense of the accuracy by using a Gaussian approximation of the *a posteriori* density function given by (3.33). A Gaussian approximation is developed in Appendix E on page 111. The approximation is predicated on a non-linear least squared error analysis and does not require integration. The resulting approximate *a posteriori* density function yields a Gaussian pdf for a pdf that is not Gaussian, but it is a reasonable approximation if the errors on the ToA estimates are small and the target is not near one of the gateways.

3.5 Finding the Probability of a Specified Credible Region

The conditional joint density function $f_{x,y|\widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y}(x, y \mid \boldsymbol{\tau}, \widehat{\boldsymbol{g}}_x, \widehat{\boldsymbol{g}}_y)$ contains all the information in $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ gained through the ToA and gateways' location measurements. In terms of the probabilistic experiment it contains all the information gained through the observations of the random variables $\widehat{\mathsf{ToA}}(\boldsymbol{\zeta})$, $\widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})$, $\widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})$. This joint density is used to obtain the probability the target is in any specified region, which is more properly called a specified credible region, of the x-y plane. That probability is the the volume under the joint density that is supported by the specified credible region.

3.6 Summary of Point and Region Estimators and Associated Terminology

Often estimates are obtained without using the joint density. Many of these estimates, especially those obtained from adhoc algorithms, cannot be obtained from the conditional joint density, but the MMSE and ML estimators can.

MMSE estimate: The MMSE estimate is the conditional mean, which is obtained by taking the first moment. An MMSE estimate of the *a posteriori* pdf reduces the information in the conditional joint density to a single value. Obviously, this is a massive compression of the information in the *a posteriori* pdf.

ML estimate: The ML estimate for a single random variable is the value of the argument that maximizes the conditional marginal density. The estimate obtained is the ML estimate if the *a priori* pdf is widely uniform. The joint ML estimates of two random variables are the two coordinates that jointly maximize the conditional joint density. In general the marginal ML estimates, which are the values that maximize the respective marginal densities, differ from the joint ML estimates. Since the joint ML estimates are pair of points the information in their values, this is a massive compression of the information in the *a posteriori* pdf.

The conditional joint density $f_{x,y|\widehat{\mathsf{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y}(x,y \mid \boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y)$ is a function of two real variables and as such contains a massive amount of information. The information is often characterized or summarized in a more condensed form. Some of these forms are listed below.

Conditional mean: The conditional mean condenses the conditional joint density to a single coordinate, which is either x_{MMSE} or y_{MMSE}. The conditional means of $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ are their MMSE estimates.

Maximum likelihood: The pair of coordinates, $(x_{\text{ML}}, y_{\text{ML}})$, that maximize the conditional joint density when $f_{x,y}(x,y)$ is widely constant is the maximum likelihood estimate of the target's location.

Conditional variance: The conditional variances of $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ measure the spread of the conditional joint density. They are given by $\mathbb{E}[(x(\boldsymbol{\zeta}) - x_{\mathrm{MMSE}})^2 \mid \widehat{\mathbf{ToA}}(\boldsymbol{\zeta}) = \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y]$ and $\mathbb{E}[(y(\boldsymbol{\zeta}) - y_{\mathrm{MMSE}})^2 \mid \widehat{\mathbf{ToA}}(\boldsymbol{\zeta}) = \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y]$.

One-dimension densities: The joint density is often summarized, again with considerable loss of information, to the density for the Euclidean distance from the MMSE estimate. That is, the joint density can be compressed to the marginal density

$$f_{\rho \left| \widehat{\mathbf{ToA}},\, \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\, \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y \right.}(\rho \mid \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y)$$

for random variable $\rho(\boldsymbol{\zeta}) = \sqrt{(x(\boldsymbol{\zeta}) - x_{\mathrm{MMSE}})^2 + (y(\boldsymbol{\zeta}) - y_{\mathrm{MMSE}})^2}$. Such a density is easily obtained with numerical integration.

Probability for a Bayes Credible Region: A *Bayes credible region*, which is often shortened to *credible region*, is any specified region in the domain of the *a posteriori* probability density function (pdf). There is a probability associated with every specified credible region. This probability is the volume under the *a posteriori* pdf that is supported by the specified Bayes credible region. For example, if a credible region was specified for the location of the target, then the probability that the target is within that region is the volume under the joint *a posteriori* pdf supported by the specified credible region.

A Bayes credible region is in some ways quite similar to a confidence region, but in other ways it is quite different. The two regions are similar in that they both have a probability associated with them. The probability associated with a Bayes credible region is the probability a parameter or set of parameters, e.g. the coordinates of the target, fall within the specified credible region. The probability associated with a confidence region is the probability that an estimate obtained from a specified estimator will fall into the confidence region.

There are two main differences:

1. The Bayes credible region is specified and the probability associated with it is

calculated where as the probability associated with a confidence region is specified and the confidence region itself is estimated.

2. The probability associated with a Bayes credible region is calculated from a single experimental outcome using the *a posteriori* pdf obtained from that outcome. The confidence region is estimated from the outcomes of a large number of experiments. The confidence region is the connected region with the smallest *area* where the number of estimates in the confidence region satisfies the relative frequency definition for the specified probability.

   For example, if the experiment was repeated N times to obtain N estimates and the specified probability was P_{CONF}, then the confidence region is the smallest connected region containing $N \times P_{\mathsf{CONF}}$ of the N estimates.

Contour plots: The credible regions defined by a *a posteriori* pdf can be constructed from a series of level curves, which are better known as contours. Each closed contour becomes a credible region and can be labeled by a probability that is the volume under $f_{x,y\left|\widehat{\mathsf{ToA}},\, \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})=\widehat{\boldsymbol{g}}_x,\, \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})=\widehat{\boldsymbol{g}}_y\right.}(x, y \mid \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y)$ that is supported by the credible region. That is, each contour is labeled with the probability that $(x(\boldsymbol{\zeta}), y(\boldsymbol{\zeta}))$ is located in the area enclosed by the contour.

It is important that the relative values of point estimators, like the MMSE, MAP and ML estimators, compared to credible regions and the joint density

$$f_{x,y\left|\widehat{\mathsf{ToA}},\, \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta})=\widehat{\boldsymbol{g}}_x,\, \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta})=\widehat{\boldsymbol{g}}_y\right.}(x, y \mid \boldsymbol{\tau},\ \widehat{\boldsymbol{g}}_x(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_x,\ \widehat{\boldsymbol{g}}_y(\boldsymbol{\zeta}) = \widehat{\boldsymbol{g}}_y)$$

be put into perspective. It may seem that the sole purpose for obtaining the joint density was to obtain the MMSE and ML estimates. Thinking along those lines would suggest that if a closed form expression for the MMSE and ML estimators were somehow obtainable, there would be no value in finding the probability associated with a specified credible region. However, that is not the case.

There is considerably more information in the *a posteriori* pdf than any point estimator or the probability associated with a particular credible region. In fact, the *a posteriori*

pdf could be used to generate point estimates or probabilities associated with any credible regions. To put the worth of the probabilities associated with credible regions and the point estimates in perspective, consider the example where the target happens to be a LoRa sensor worn by a missing child. A point estimator would provide a good starting point for the search, but the contour lines defining credible regions and the associated probabilities would be great value in organizing a grid search for the child.

4. Corroboration and Utility of the *a Posteriori* PDF and Multiple Transmissions

4.1 Corroboration of (3.33) in the Absence of Gateway Location Error

Unfortunately, for reasons that will be explained later, it is not possible to corroborate the accuracy of (3.33) with a Monte Carlo simulation. Instead, an example that compares the *a posteriori* pdf given by (3.33) to the rough Gaussian approximation obtained using Foy's linearized model, introduced in Appendix E on page 111, is used to corroborate the development of (3.33). In this section it is assumed the locations of the gateways are known exactly, which means the exact values are used for the observations in (3.33) and $\sigma^2_{t_\ell}$ is used in place of $\sigma^2_{\mathsf{sum}_\ell}$. The example is based on a system of 4 gateways positioned as shown in Fig. 4.1 with gateways 1 to 4 having co-ordinates $(4, 4)$ km, $(-4, 4)$ km, $(-4, -4)$ km, and $(4, -4)$ km, respectively, is used for that purpose. The target is located at the origin in the same horizontal plane as the gateways. The gateway spacing and location of the target were chosen so that all the range equations are nearly linear. This was done to ensure good agreement between *a posteriori* pdf and the Gaussian approximation. The measurement errors on the times-of-arrival at the four gateways are independent with identical zero-mean Gaussian distributions that have standard deviations of 25 ns.

The mean of the joint Gaussian approximation, i.e., $(\bar{x}, \bar{y})$, is taken to be the LS solution to the L non-linear equations given by (2.3). The covariance matrix for $x(\boldsymbol{\zeta})$ and $y(\boldsymbol{\zeta})$ is the 2×2 upper left block of the covariance matrix given by (2.15).

Fig. 4.2 shows two marginal *a posteriori* pdfs, both for the x-coordinate of the target,

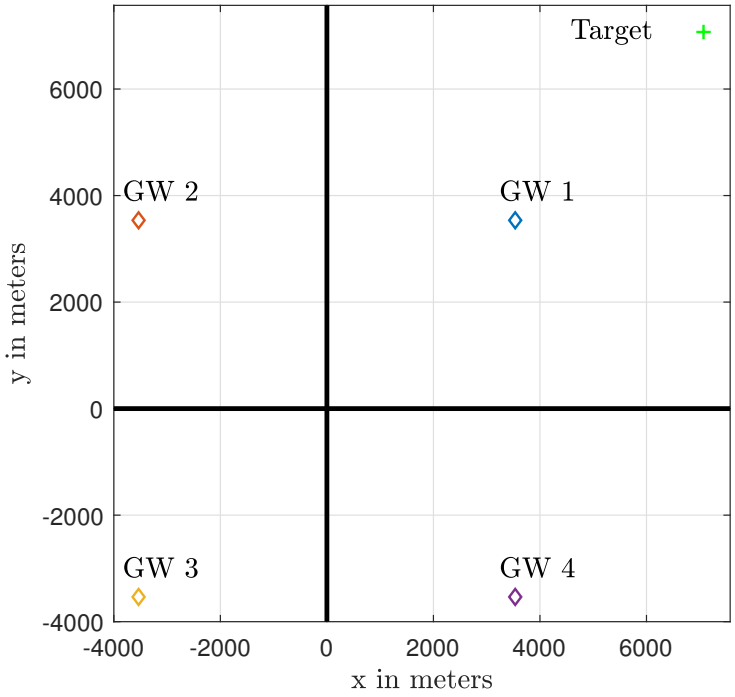

Figure 4.1: Geometry for the experiment.

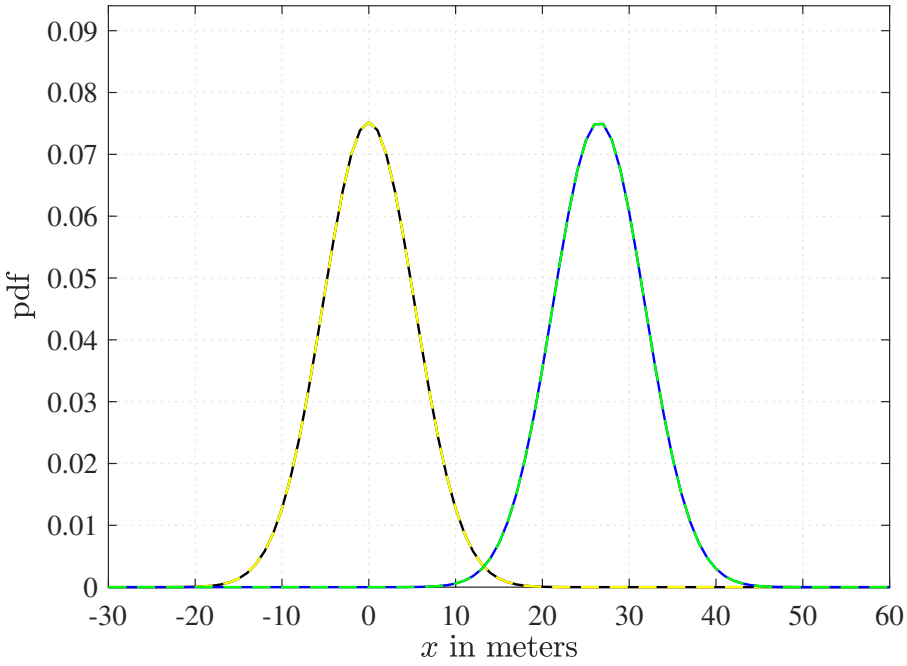

Figure 4.2: Corroboration of the validity of (3.33) (solid black and solid blue) by comparison with the Gaussian approximations (dashed yellow and dashed green).

plotted in solid black and solid blue lines. The pdfs are then overwritten with their Gaussian approximations which are the dashed yellow and green lines. The two *a posteriori* pdfs were

generated from (3.33) using a widely uniform *a priori* pdf for two different outcomes: one would be considered a lucky outcome and the other a "statistical outlier". The "lucky" outcome has no errors on the times-of-arrival at gateways 1, 2, 3 and 4 and the "statistical outlier" outcome has errors of $-125\,\text{ns}$, $0\,\text{ns}$, $125\,\text{ns}$ and $0\,\text{ns}$, respectively. As explained earlier, it is assumed the gateways' locations are known so error in the gateways' location $(0, 0)$ for both the "lucky" and "statistical outlier". The "lucky" outcome produced the solid black (*a posteriori*) and dashed yellow (Gaussian approximation) curves. Both curves are centered at the origin, which is expected for an error free outcome. Clearly, there is very good agreement between the *a posteriori* pdf given by (3.33) and approximation-based Gaussian pdf as the dashed yellow curve seems to coincide with the solid black curve. The "statistical outlier" outcome generated the curves centered at about 27 m. Obviously, the measurement error in the ToAs causes the shift. Again, there is very good agreement between the *a posteriori* pdf and the approximation to it. This suggests that the equation for the *a posteriori* pdf given by (3.33) is correct for the case where the locations of the gateways are known.

4.2 Value of (3.33) in the Absence of Gateway Location Error

The times-of-flight are non-linear functions of the target's location. Therefore, if the times-of-arrival have a Gaussian distribution, then the *a posteriori* pdf of the target's location cannot be Gaussian. The geometry is such that the equation for time-of-flight becomes increasingly non-linear as the distance between the target and gateway decreases. This means the Gaussian approximation becomes increasingly inaccurate as the distances between the target and gateways decrease. Therefore, the effects of the non-linearities are more easily demonstrated in a system where the gateways are close together. For this purpose the gateways are moved closer together to be located on a square like that in Fig. 4.1, except with coordinates $(\pm 100, \pm 100)\,\text{m}$. The target is positioned at $(85, 85)\,\text{m}$, which places it $21.2\,\text{m}$ from gateway 1. This is a reasonable scenario in an urban setting.

Two variances are considered for the error on the ToA measurements: one optimistic, the other more realistic. Both distributions are zero-mean Gaussian, but the optimistic

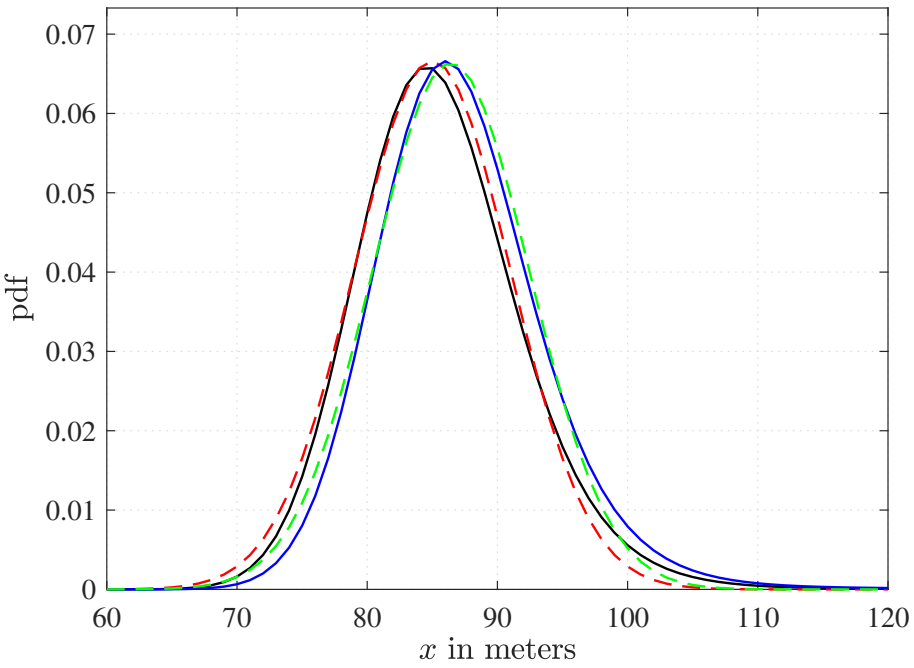

Figure 4.3: Marginal densities for the x co-ordinate of the target for two sets of measurements when the standard deviation for the error on the times-of-arrival is 25 ns.

distribution has a standard deviation of 25 ns, which translates to 7.5 m in ToF distance, and the more realistic distribution has a standard deviation of 50 ns, which translates to 15 m in ToF distance. Two observations are drawn from each distribution. One is the "lucky" observation, which has no measurement error, and the other is the more typical observation, which has measurement error of plus one standard deviation on gateways 2 and 4.

Marginal *a posteriori* pdfs and their Gaussian approximations for the optimistic distribution of ToA measurement error that has $\sigma_{t_\ell} = 25$ns, are shown in Fig. 4.3. The solid lines show the true *a posteriori* densities for the x-coordinate of the target's location. The dashed lines show the Gaussian approximations. The solid black and the dashed red curves result from the "lucky" observation (no ToA measurement error), whereas the solid blue and the dashed green curves result from the "more typical" observation (some ToA measurement error). Comparing the dashed curves, which are Gaussian pdfs, to the solid lines it is clear the *a posteriori* pdfs are not Gaussian, but the Gaussian approximations are reasonable.

Fig. 4.4 shows the marginal *a posteriori* pdfs and their Gaussian approximations for the more realistic distributions for ToA measurement error, which has $\sigma_{t_\ell} = 50$ ns. Similar to

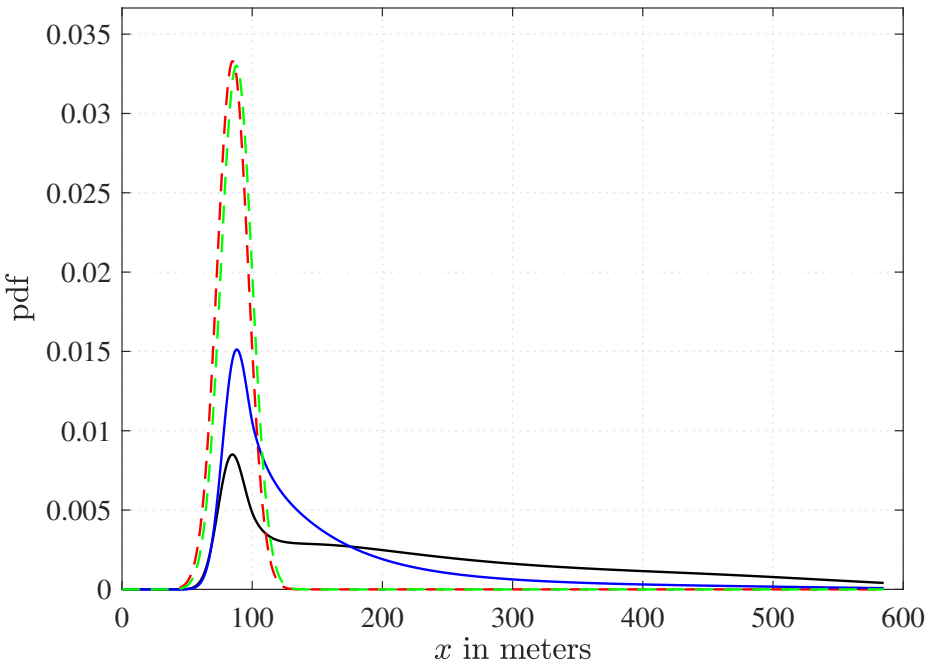

Figure 4.4: Marginal densities for the x co-ordinate of the target for two sets of measurements when the standard deviation for the error on the times-of-arrival is 50 ns.

Fig. 4.3, the solid black and the dashed red curves are the *a posteriori* pdf and the Gaussian approximation, respectively, resulting from the lucky observation. The solid blue and the dashed green curves result from the more typical observation. It is very clear that the marginal *a posteriori* pdfs are not at all Gaussian and that they should not be modeled as such. A non-Gaussian *a posteriori* pdf indicates at least one of the time-of-flight equations is significantly non-linear over the reach of measurement error.

It is important to mention that the Gaussian approximation is not always easy to obtain. With the gateways located 200m apart one of the ToF equations becomes so non-linear when the target is close to a gateway the iterative search for the nonlinear LS solution has stability issues. For example, when the target was 1 meter from a gateway the steepest decent algorithm had to be throttled with a convergence factor of 0.001 for it to converge.

Upon noticing the two *a posteriori* pdfs in Figure 4.4 are far from Gaussian, one may wonder why these particular two pdfs were not corroborated. Unfortunately, all efforts to find a method of corroboration are failed. One may think that (3.33) should be verifiable by Monte Carlo simulations in all scenarios, but that is not the case. Each *a posteriori* pdf

is computed entirely from a single set of ToA measurements and is unique to the errors in that set of measurements. There is no way to "roll some dice" and then compute individual points on the surface of the joint *a posteriori* pdf or on the curve of a marginal pdf. It was for this reason that (3.33) was corroborated using a system and set of parameters such that the ToF equations could be accurately approximated with Foy's linearized model.

Finally, a contrived scenario where the *a priori* pdf is something other than uniform. Suppose a target in inoperable in water and there is a large body of water in range of the gateways. Further, suppose the shoreline is parallel to the y-axis, but its intercept with the x-axis is not known exactly due to variation in water level.

For example, suppose the gateways are located in a horizontal plane as shown in Fig. 4.1, while the target is located in the same horizontal plane at $x = y = 7070\,\text{m}$ and suppose the x-axis intercept for the shoreline is uniform $(7100, 7150)$ m. Then the *a priori* pdf that accounts for the uncertainty in the x-intercept of the shoreline could be reasonably represented by[‡]

$$f_{x,y}(x,y) = \begin{cases} \epsilon, & -50000 < x < 7100 \text{ m} \\ \frac{\epsilon \times (7150 - x)}{50}, & 7100 \leq x \leq 7150 \text{ m} \\ 0, & x > 7150 \text{ m}, \end{cases} \tag{4.1}$$

where ϵ is the constant that makes the volume under $f_{x,y}(x,y)$ equal to 1. Normally the constant ϵ would have to be very small to make $f_{x,y}(x,y)$ a proper joint density. However, $f_{x,y}(x,y)$ appears in both the numerator and denominator of (3.33), so it can be scaled without changing the result. Therefore, for the purposes of computing the left side of (3.33), ϵ can be assigned any value, the most logical of which is 1.

Using the *a priori* pdf given by (4.1), the system of gateways in Fig. 4.1, the realistic value of 50 ns for σ_{t_ℓ} and the lucky observation (no measurement error on the ToAs), Eqn. (3.33) yields the joint *a posteriori* pdf shown in Fig. 4.5. Contour lines on the joint pdf shown in Fig. 4.5 provide credible regions. Two credible regions, as well as the MMSE and MAP

[‡]This *a priori* pdf is chosen to reasonably represent the shoreline in the contrived example. In practice, data could be collected from, say, satellite imagery and a *a priori* pdf is then generated from the data.

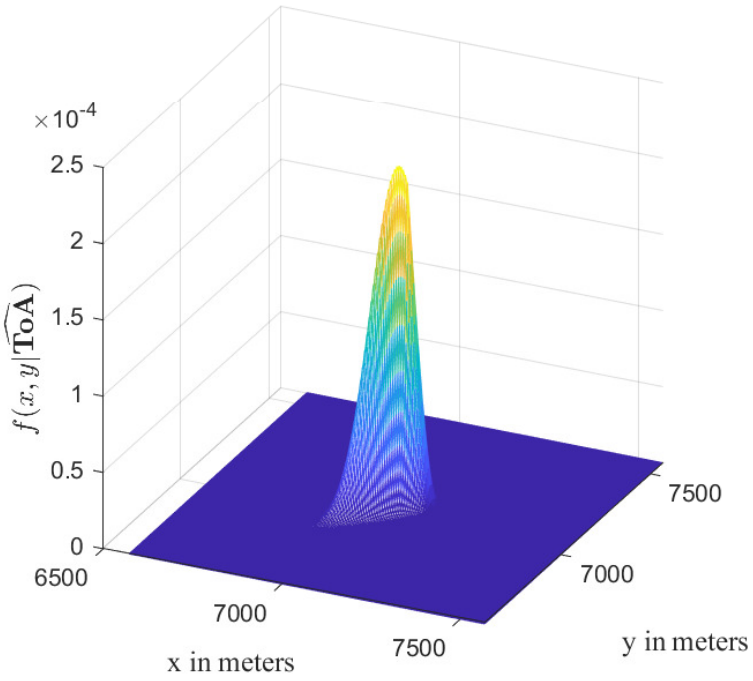

Figure 4.5: The *a posteriori* density when the target is in the vicinity of a body of water.

estimates, are shown in Fig. 4.6. The asymmetry of the contour lines indicate that the *a priori* density has changed the *a posteriori* pdf and that the pdf is not jointly Gaussian. This is further verified in plots in Fig. 4.7 that show the marginal density $f_{x|\widehat{\mathbf{ToA}}}(x|\boldsymbol{\tau})$, which is the marginal density for the x-coordinate of the target, along with the associated Gaussian approximation.

Asymmetry in the *a posteriori* pdf causes separation between the MAP and MMSE estimates, which means the MAP estimate cannot serve as a proxy for the MMSE estimate. In this particular case the *a priori* density that models the presence of the body of water did not change the MAP estimate, but did change the MMSE estimate. The MAP estimate remains identical to the ML estimate since the *a priori* density is flat where $\mathrm{e}^{-\frac{\gamma}{2}}$ peaks. Clearly an *a priori* pdf can significantly change the MMSE estimate, making the probabilistic approach superior to LS-based estimators when *a priori* information is available.

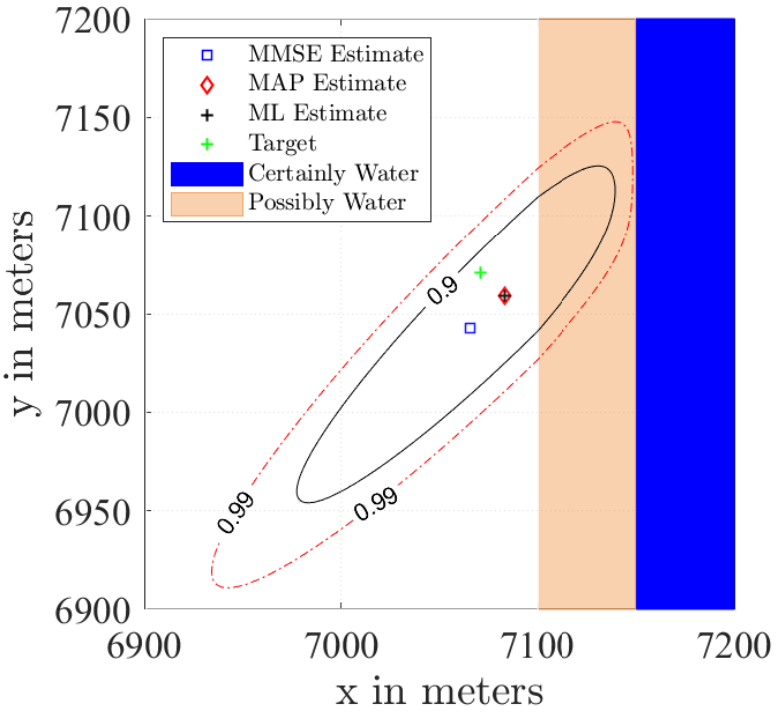

Figure 4.6: Two credible regions and point estimates of the target when it is in the vicinity of a body of water.

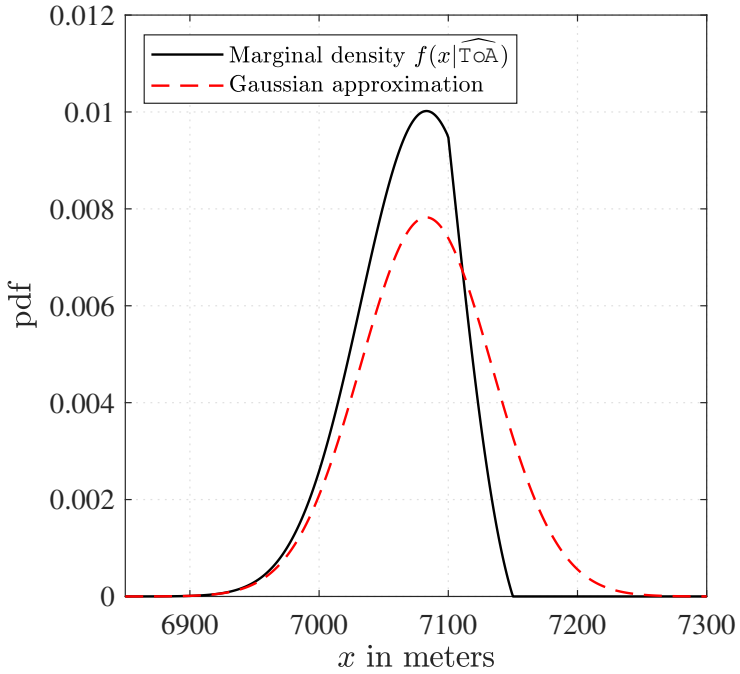

Figure 4.7: The marginal density $f_{x|\widehat{\mathbf{ToA}}}(x|\boldsymbol{\tau})$ when the target is in the vicinity of a body of water.

4.3 Effects of Errors in Gateways' Measured Locations

Uncertainty in the locations of the gateways contributes to the *a posteriori* pdf in two ways:

1. The variance of error contributes to the spread, or the "size of the footprint".
2. The measurement error, i.e., the error in the observations, causes a positional shift and also contributes to the spread.

The effect of the error in the observation can be explained using (3.33) on page 47, by letting $\sigma_{x_\ell}^2 \to 0$ and $\sigma_{y_\ell}^2 \to 0$ for $\ell = 1, 2, \ldots, L$. With $\sigma_{x_\ell}^2$ and $\sigma_{y_\ell}^2$ set to zero the gateways' locations are treated as known. This means if there was an error in an observation it would change the system geometry, causing the maximum *a posteriori* (MAP) estimate to change. Causing the MAP estimates to change, in essence, is equivalent to shifting the pdf, since the MAP estimates coincide with the peak of the *a posteriori* pdf.

Furthermore, the error in the gateways' locations will change the size of the aperture (receiving cross section or effective area) of the system of gateways as seen by the target. Since the accuracy of localization algorithms is somewhat proportional to the size of the aperture, one would expect the relative change in the spread of the pdf to be somewhat proportional to the relative change in aperture. Therefore, errors in the measurements of the gateways' locations will affect the spread of the *a posteriori* pdf with the effect being more significant when the aperture is small.

The variances $\sigma_{x_\ell}^2$ and $\sigma_{y_\ell}^2$ affect the spread and shape of the *a posteriori* pdf. Under the special condition where $\sigma_{x_\ell} = \sigma_{y_\ell}$ and neither σ_{x_ℓ} nor σ_{t_ℓ} depend on ℓ, it can be shown that the ML estimate, i.e., the MAP estimate for a widely uniform *a priori* pdf, does not depend on either σ_{x_ℓ} or σ_{t_ℓ}. While the authors cannot provide a mathematical proof for the general case, simulations of cases where $\sigma_{x_\ell}^2 = \sigma_{y_\ell}^2$ and both $\sigma_{x_\ell}^2$ and $\sigma_{t_\ell}^2$ change with ℓ produced the same ML estimate within the accuracy of computation. This suggests that under conditions where $\sigma_{x_\ell} = \sigma_{y_\ell}$, the ML estimate depends only on the error in the observations, i.e., measurements of the gateways' locations. This means $\sigma_{x_\ell}^2$ and $\sigma_{y_\ell}^2$ will affect on the shape and spread of

the *a posteriori* pdf, but will not shift it.

The dependencies discussed above can be illustrated by plotting the critical regions for four scenarios. Consider a system with $L = 4$ gateways located at $\left(\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{2}}\right)$ km, $\left(-\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{2}}\right)$ km, $\left(-\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$ km and $\left(\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$ km. Suppose the ToAs are measured without error, i.e., $\sigma_{t_\ell} \to 0$ and $\boldsymbol{\tau}_\ell$ has no error. Now consider four combinations of variances of error and measurement errors for the gateways:

Case 1: $\sigma_{x_\ell} = \sigma_{y_\ell} = 10\text{m}\ \forall \ell$, gateway measurement errors are $(0, 0)\ \forall \ell$.

Case 2: $\sigma_{x_\ell} = \sigma_{y_\ell} = 10\text{m}\ \forall \ell$, gateway measurement errors are $(-5, 0)$m, $(0, -5)$m, $(5, 0)$m and $(0, 5)$m.

Case 3: $\sigma_{x_\ell} = \sigma_{y_\ell} = 30\text{m}\ \forall \ell$, gateway measurement errors are $(0, 0)\ \forall \ell$.

Case 4: $\sigma_{x_\ell} = \sigma_{y_\ell} = 30\text{m}\ \forall \ell$, gateway measurement errors are $(5, 0)$m, $(0, 5)$m, $(-5, 0)$m and $(0, -5)$m.

Note the errors in Case 4 have opposite signs to the errors in Case 2. The contour lines on the *a posteriori* pdf that encompass a probability of 0.7 for the four cases above are shown in Fig. 4.8. The solid black and solid green lines are the plots for Case 1 and Case 2, respectively. The dashed red line and dashed blue lines are plots of Cases 3 and 4, respectively. All four curves appear to be somewhat elliptical so they will loosely referred to as elliptical. The "elliptical curves" for Cases 1 and 2 show a slight change in a position, but no significant change in shape. The curves for Cases 3 and 4 indicate the same thing, except the change in position is in the opposite direction. One can conclude that errors in the measurements of the gateways' locations affect the position of the critical region, but since the aperture is quite large and the errors are relatively small the errors do not significantly change the size of the aperture. In this case the errors in the measurements of the gateway's location will have very little effect on the size and shape of the critical region.

Comparison of the critical regions for Cases 1 and 3 suggests the size and shape of the critical region depend on the variance of error on the measurement of the gateways' locations. The same conclusion is reached by comparing the critical regions for Cases 2 and 4.

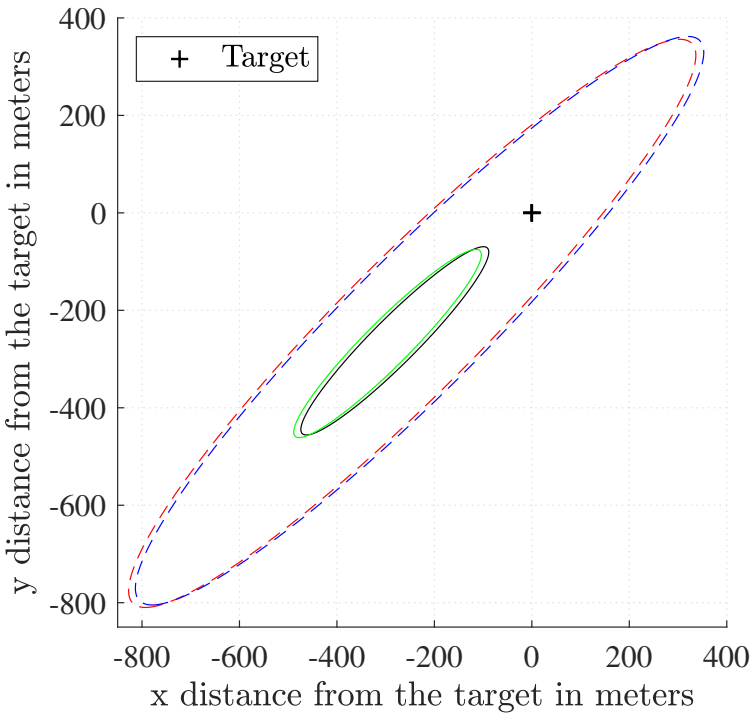

Figure 4.8: Critical regions with probability 0.70 for different measurements and variances of the gateways' locations.

4.4 Multiple Transmissions

The *a posteriori* pdf given by (3.33) was developed based on a single transmission of a message from the target. From that transmission one set ToA measurements and one set of gateway location measurements were made and those measurements were used in (3.33). Should the target transmit multiple times from a fixed location, the *a posteriori* pdf can be improved after each transmission. Of course the ToAs must be measured at each gateway for each transmission. However, when it comes to the measurement of the gateways' locations, there are two different cases to consider:

1. The locations of the gateways are measured each time a transmission is received from the target. For example, this case applies to the scenario where gateways are mobile.

2. The locations of the gateways are measured just once. This applies if the locations of the gateways are fixed.

In the first case, the *a posteriori* pdf can be updated iteratively after each transmission.

Specifically, since the measurements made for each transmission are independent, (3.33) can be applied iteratively with the *a priori* pdf being the *a posteriori* pdf calculated in the previous iteration.

In the second case, (3.33) cannot be applied iteratively since the measured gateway coordinates are not independent from one transmission to the next. A rigorous analysis of this case may be possible, but it is believed the analysis would be very tedious and likely produce an expression that is too complicated to be of great practical value. Instead it is proposed that the arrival times of the transmissions be averaged and that average be used in (3.33) as a single measurement. Of course the averaged ToA measurements would have a smaller variance and the variances σ_{t_ℓ} in (3.33) would have to be replaced with the smaller variances. This approach does not use all of the information in the measured ToAs from the multiple transmissions so (3.33) will produce a pdf that is a pessimistic approximation to the true *a posteriori* pdf.

The validity of using averages for the time-of-arrival is demonstrated with the following mathematical argument below.

Let the target make N transmissions and let the times the N transmissions arrive at Gateway ℓ be considered independent random variables. Also, let the RVs associated with the nth transmission be defined on sample space $\boldsymbol{S}^{(n)} = \{\zeta_1^{(n)}, \zeta_2^{(n)}, \dots\}$. Furthermore, let $\widehat{\mathsf{ToA}}_\ell^{(n)}(\zeta^{(n)})$, or $\boldsymbol{\tau}_\ell^{(n)}$ for short, symbolize the measured time of arrival of the nth transmission at Gateway ℓ. Then the average of the N times of arrival at Gateway ℓ is given by

$$\begin{aligned}
\overline{\boldsymbol{\tau}}_\ell &= \frac{1}{N}\sum_{n=1}^{N} \boldsymbol{\tau}_\ell^{(n)} \\
&= \frac{1}{N}\sum_{n=1}^{N} \left(t^{(n)}(\zeta^{(n)}) + \mathsf{ToF}_\ell(\zeta^{(n)}) + \Delta T_\ell^{(n)}(\zeta^{(n)}) \right) \\
&= \bar{t} + \mathsf{ToF}_\ell + \overline{\Delta T}_\ell. \qquad (4.2)
\end{aligned}$$

where $t^{(n)}(\zeta^{(n)})$ is the time of nth transmission, $\bar{t} = \frac{1}{N}\sum_{n=1}^{N} t^{(n)}(\zeta^{(n)})$ is the average of N transmission times, $\mathsf{ToF}_\ell(\zeta^{(1)}) = \mathsf{ToF}_\ell(\zeta^{(2)}) = \cdots = \mathsf{ToF}_\ell(\zeta^{(N)}) = \mathsf{ToF}_\ell$ is the time of flight

from the target to gateway ℓ (which is the same for every transmission), and $\Delta T_\ell^{(n)}(\zeta^{(n)})$ is the measurement error on the time the nth transmission arrives at gateway ℓ and $\overline{\Delta T}_\ell = \frac{1}{N}\sum_{n=1}^{N} \Delta T_\ell^{(n)}(\zeta^{(n)})$.

Note that $\overline{\Delta T}_\ell$ is an observation of a random variable that is the sum of N independent zero-mean Gaussian RVs, each with variance $\frac{(\sigma_{t_\ell}^{(n)})^2}{N^2}$. Such a random variable will be a zero-mean Gaussian RV with variance

$$\overline{\sigma}_{t_\ell}^2 = \frac{1}{N^2}\sum_{n=1}^{N}(\sigma_{t_\ell}^{(n)})^2.$$

Since (3.33) is invariant to the time of transmission, (4.2) has the form of a single transmission and therefore can be used as such.

It is pointed out that, although the target and gateways are fixed, it is possible for $\sigma_{t_\ell}^{(n)}$ to depend on the time of arrival, i.e., depend on n. This would be the case if a transmission from another target collided or partially collided with the arrival of a transmission. This could even cause the arrival of the nth transmission to go undetected. Such a case will be discussed more later. For now it is assumed $\sigma_{t_\ell}^{(n)}$ does not depend on n. Then the variance of the average of the times of arrival at gateway ℓ is

$$\overline{\sigma}_{t_\ell}^2 = \frac{1}{N}\sigma_{t_\ell}^2 \quad \text{for } \sigma_{t_\ell}^{(n)} = \sigma_{t_\ell};\ n = 1, 2, \ldots, N. \tag{4.3}$$

Then (3.33) can be used to find an approximate *a posteriori* pdf for multiple transmissions from a fixed target to gateways whose locations are fixed and have been measured only once. This is done by substituting $\overline{\sigma}_{t_\ell}^2 = \frac{\sigma_{t_\ell}^2}{N}$ for $\sigma_{t_\ell}^2$ in (3.37), which modifies $\sigma_{\mathsf{sum}_\ell}^2$, then substituting the modified $\sigma_{\mathsf{sum}_\ell}^2$ into (3.36) as well as $\overline{\tau}_\ell$ for τ_ℓ into (3.36), which modifies γ, then using the modified γ in (3.33) to get an approximate *a posteriori* pdf.

While the resulting pdf will be an improvement over the *a posteriori* pdf for a single transmission, it will not be the true *a posteriori* for N transmissions since it can be shown that the *a posteriori* pdf for N transmissions cannot be reduced to a function of $\sum_{n=1}^{N} \tau_\ell^{(n)}$ for $\ell = 1, \ldots, L$.

4.5 Effects of Multiple Transmissions

Recall from Section 4.4 that there are two scenarios for multiple transmissions: (1) The locations of the gateways are measured each time a transmission is received from the target, and (2) The locations of the gateways are measured just once. The quality of an approximate *a posteriori* pdf developed using multiple transmissions for the second scenario can be assessed experimentally using the first scenario. In the first scenario, (3.33) is applied iteratively, once for each transmission, to produce the true *a posteriori* pdf. To verify the quality of the pdf resulting from using the average $\overline{\tau}_\ell$ for τ_ℓ in (3.33), the average values of the locations of the gateways must also be used for $\widehat{x}_\ell$ and $\widehat{y}_\ell$. These average values are given by

$$\widehat{x}_\ell^{(\mathsf{avg})} = \frac{1}{N}\sum_{n=1}^{N} \widehat{x}_\ell^{(n)}, \qquad \widehat{y}_\ell^{(\mathsf{avg})} = \frac{1}{N}\sum_{n=1}^{N} \widehat{y}_\ell^{(n)}.$$

The variances of random variables $\widehat{x}_\ell^{(\mathsf{avg})}(\boldsymbol{\zeta})$ and $\widehat{y}_\ell^{(\mathsf{avg})}(\boldsymbol{\zeta})$, which are

$$\overline{\sigma}_{x_\ell}^2 = \frac{1}{N^2}\sum_{n=1}^{N} \left(\sigma_{x_\ell}^{(n)}\right)^2$$

and

$$\overline{\sigma}_{y_\ell}^2 = \frac{1}{N^2}\sum_{n=1}^{N} \left(\sigma_{y_\ell}^{(n)}\right)^2,$$

respectively, must be used in (3.33) for $\sigma_{x_\ell}^2$ and $\sigma_{y_\ell}^2$. Should $\sigma_{x_\ell}^{(n)} = \sigma_{x_\ell}$ and $\sigma_{y_\ell}^{(n)} = \sigma_{y_\ell}$ for $n = 1, 2, \ldots, N$, then $\overline{\sigma}_{x_\ell}^2 = \frac{1}{N}\sigma_{x_\ell}^2$ and $\overline{\sigma}_{y_\ell}^2 = \frac{1}{N}\sigma_{y_\ell}^2$.

The pdf computed using the average values $\overline{\tau}_\ell$, $\widehat{x}_\ell^{<avg>}$ and $\widehat{y}_\ell^{<avg>}$ as well as variances $\overline{\sigma}_{t_\ell}^2$, $\overline{\sigma}_{x_\ell}^2$ and $\overline{\sigma}_{y_\ell}^2$ in (3.33) can be compared to the true *a posteriori* pdf computed by applying (3.33) iteratively using $\tau_\ell^{(n)}$, $x_\ell^{(n)}$, $y_\ell^{(n)}$, $(\sigma_{t_\ell}^{(n)})^2$, $(\sigma_{x_\ell}^{(n)})^2$ and $(\sigma_{y_\ell}^{(n)})^2$. The similarity of the two pdfs can be established by comparing critical regions defined by the contour lines on the pdfs. Two such critical regions will be compared: one yielding a probability of 0.7 and the other yielding a probability of 0.99. The system geometry used for the comparison has $L = 4$ gateways located at $\left(\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{2}}\right)$ km, $\left(-\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{2}}\right)$ km, $\left(-\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$ km and $\left(\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$ km, respectively, and the target is located at $\left(\frac{10}{\sqrt{2}}, \frac{10}{\sqrt{2}}\right)$ km.

The comparisons are made for $\sigma_{t_\ell} = 100$ ns and $\sigma_{x_\ell} = \sigma_{y_\ell} = 30$ m for $\ell = 1, 2, \ldots, L$ and for $N = 10$, $N = 33$ and $N = 100$ for critical regions with probabilities 0.7 and 0.99.

The contours that encompass critical regions with probability 0.7 for $N = 10$, 33 and 100 are plotted in Fig. 4.9. The quantities $\overline{\tau}_\ell$, $\widehat{x}_\ell^{<\mathsf{avg}>}$ and $\widehat{y}_\ell^{<\mathsf{avg}>}$ are the averages of samples drawn from independent Gaussian random variables with means $\widehat{\mathsf{ToA}}_\ell(\zeta)$, x_ℓ and y_ℓ and variances $\sigma_{t_\ell}^2$, $\sigma_{x_\ell}^2$ and $\sigma_{y_\ell}^2$. The agreement between the true *a posteriori* pdf (applying (3.33) iteratively) and the approximation (applying (3.33) once using $\overline{\tau}_\ell$, $\widehat{x}_\ell^{<\mathsf{avg}>}$ and $\widehat{y}_\ell^{<\mathsf{avg}>}$) is so good that there appears to be just three somewhat elliptical curves. The true *a posteriori* curves were plotted first in solid red and the approximations were plotted after the solid curves in dashed green. The agreement is so good the dashed green curves seem to coincide with the solid red curves. The largest "elliptical" curve is for 10 transmissions and the smallest is for 100 transmissions.

The curves for contours that encompass critical regions with the probability 0.99 are shown in Fig. 4.10. Again there is very good agreement between the true and approximate *a posteriori* pdfs. Such good agreement in both of these critical regions suggests using (3.33) once with $\overline{\tau}$ for τ_ℓ, $\widehat{x}_\ell^{<\mathsf{avg}>}$ for $\widehat{x}_\ell$ and $\widehat{y}_\ell^{<\mathsf{avg}>}$ for $\widehat{y}_\ell$ will provide a reasonable approximation to the true *a posteriori* pdf for multiple transmissions from a fixed target.

It is pointed out that the average value of the ToA must be a simple average. If a weighted average is used, the same weighting must be applied to the corresponding set of the times of arrival at all L gateways. Otherwise the average transmission time, $\bar{t}$, will not be the same for all the gateways. For the same reason, if a transmission is missed at one gateway, the corresponding ToA measurements at all other gateways must be discarded.

The effect of imperfect gateway location measurements on the *a posteriori* pdf for the target's location in the second scenario, where the gateways' location are measured once, shall be illustrated with an example. The example is contrived to be a clinical version of a practical application. The gateways and target will be located as per the previous example, which is as follows: gateways 1, 2, 3 and 4 are located at $\left(\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{2}}\right)$ km, $\left(-\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{2}}\right)$ km, $\left(-\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$ km and $\left(\frac{5}{\sqrt{2}}, -\frac{5}{\sqrt{2}}\right)$ km, respectively, and the target is located at $\left(\frac{10}{\sqrt{2}}, \frac{10}{\sqrt{2}}\right)$ km.

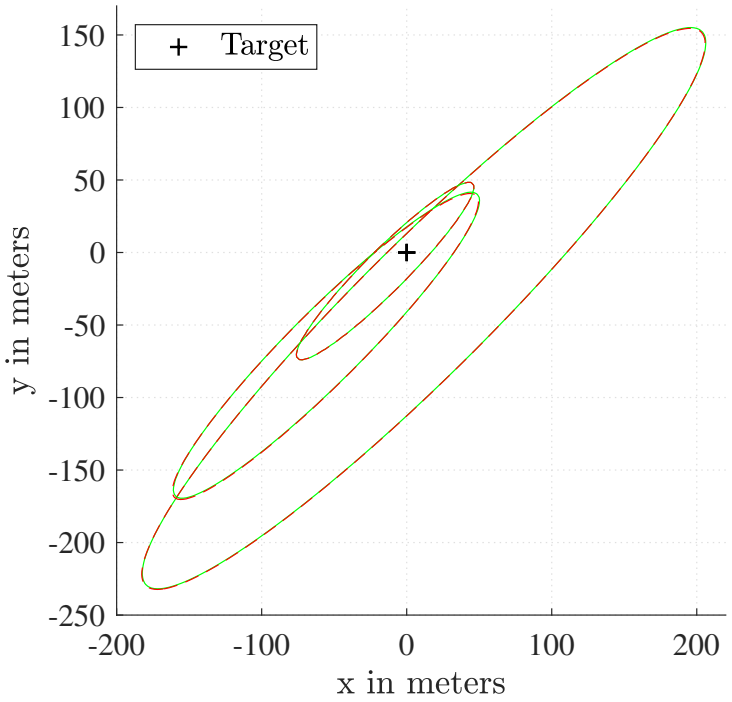

Figure 4.9: Critical regions with probability 0.70 for 10, 33 and 100 transmissions.

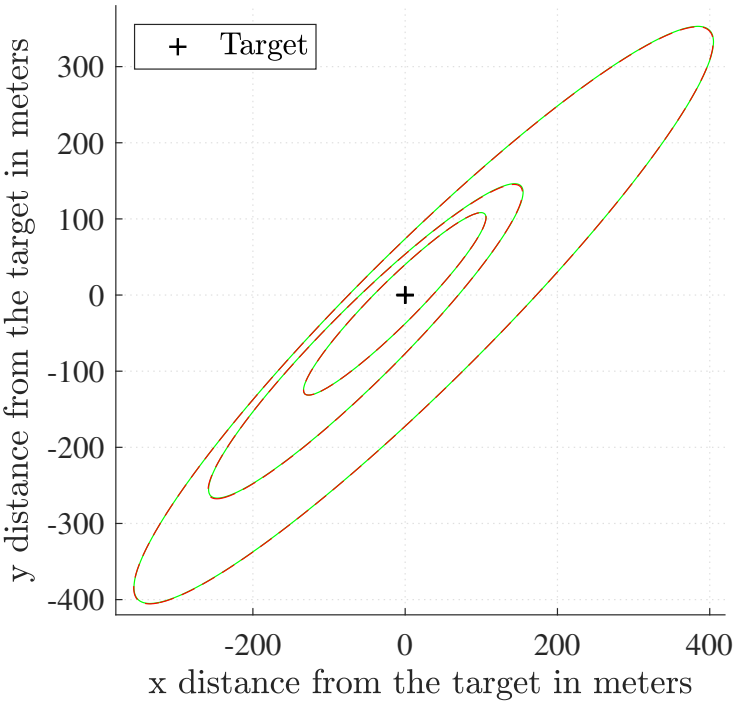

Figure 4.10: Critical regions with probabilities 0.99 for 10, 33 and 100 transmissions.

The effect of *a priori* information on the *a posteriori* pdf is illustrated by locating the target near a body of water, with the understanding that the target is inoperable in water.

The shoreline is known to be parallel to the y-axis but, perhaps due to tides or seasonal rainfall, the shoreline is not known precisely and is described by a pdf, which in this example is assumed to be uniform between $x = 7100$m and $x = 7150$m. Then the *a priori* pdf is

$$f_{\mathsf{apriori}}(x,y) = \begin{cases} \epsilon, & -50,000 < x < 7,100 \text{ m} \\ \frac{\epsilon \times (7,150-x)}{50}, & 7,100 \leq x \leq 7,150 \text{ m} \\ 0, & x > 7,150 \text{ m}, \end{cases} \tag{4.4}$$

where ϵ is the constant that makes the volume under $f_{\mathsf{apriori}}(x,y)$ equal to one.

The approximate *a posteriori* pdfs were generated with (3.33) using only measurements of the gateways' coordinates for $\widehat{x}_\ell$ and $\widehat{y}_\ell$ and the average $\overline{\tau}_\ell$ in place of τ_ℓ. The variances of the single measurements of the gateways' coordinates were used for $\sigma^2_{x_\ell}$, $\sigma^2_{y_\ell}$, but the averages $\overline{\sigma}^2_{t_\ell} = \frac{\sigma^2_{t_\ell}}{N}$ were used in place of $\sigma^2_{t_\ell}$. Then the contours that encompass a probability of 0.99 were determined. The contours were generated for four sets of repeated transmissions. The numbers of transmissions in the four sets are $N = 1$, $N = 16$, $N = 49$ and $N = 30,000$. The errors in the measurements of the gateway locations for gateways 1, 2, 3 and 4, were not chosen at random. They were either all $(0,0)$ or $(-5,0)$ m, $(0,-5)$ m, $(5,0)$ m and $(0,5)$ m, respectively, for every transmission. All except the first measured times of arrival were chosen at random from a Gaussian distribution. All L times of arrival for $n = 1$ were chosen to be error free.

Fig. 4.11 shows 8 somewhat elliptical curves, each representing a critical region. The solid lines are for no measurement error on the gateways' location and the dashed lines are for the measurement error mentioned above. The variances are $\left(\sigma^{(n)}_{t_\ell}\right)^2 = \sigma^2_{t_\ell} = (100 \text{ ns})^2$ and $\sigma^2_{x_\ell} = \sigma^2_{y_\ell} = (10 \text{ m})^2$ for $\ell = 1, 2, \ldots, L$ for all 8 curves. Since averaging reduces the effect of both the ToA variances and ToA observations, the size of the critical regions diminish with N. The largest pair of critical regions are for $N = 1$ and the smallest pair are for N=30,000. For very large N the *a posteriori* pdf will only depend on the variances $\sigma^2_{x_\ell}$ and $\sigma^2_{y_\ell}$ and the measured values of the gateways' locations. These effects of averaging are illustrated in Fig. 4.11. At $N = 1$ the critical region is largest and gets progressively smaller

as N increases. The critical region seems to converge to the smallest curve, which is the curve for N=30,000. This implies the size of the area of the smallest critical region, i.e., the critical region for $N = 30,000$, is due entirely to the variances of error in the gateways' locations. The difference between the dashed and solid lines is caused by measurement error in the gateways' locations. In this case the error causes the critical region to shift without significantly affecting the shape.

Since the area in the $N = 30,000$ critical region in Fig. 4.11 is much smaller than the $N = 1$ critical region, the size of the area of the $N = 1$ critical area is largely determined by the variances of the error in the times of arrival.

The determining factors for the size of the $N = 1$ critical region are corroborated by plotting the critical region for different values of $\sigma_{x_\ell}^2$ and $\sigma_{y_\ell}^2$ with no error on either the time of arrival or gateway location measurements. The critical region encompassed by the black curve in Fig. 4.12 is that of Fig. 4.11 for $N = 1$ and no measurement error. The critical region encompassed by the solid red line is for the same parameters except $\sigma_{x_\ell} = \sigma_{y_\ell}$ is changed from 10m to 0m. Clearly, decreasing the variance of the error in the gateways' locations decreased the size of the critical region, but not that much. The size of the critical region for $N = 1$ is more sensitive to the variances of the times of arrivals. The blue curve in Fig. 4.12 was generated with the standard deviations for all times of arrival cut in half, this is $\sigma_{t_\ell}^{(n)} = \sigma_{t_\ell} = 50$ns. The critical region encompassed by the blue curve is much smaller than the critical region encompassed by the red curve, which indicates the size of the critical region is more sensitive to the variances of the times of arrival, at least in this particular system with $\sigma_{x_\ell} = \sigma_{y_\ell} = 10$m for all ℓ and $\sigma_{t_\ell}^{(n)} = 100$ns for all ℓ and n.

The factors that determine the size of the critical region for $N = 30,000$ are corroborated in the same way. The critical region for $N = 30,000$ encompassed by the solid black curve in Fig. 4.11 is plotted again in Fig. 4.13 with the solid black curve. The dashed red curve in Fig. 4.13 encompasses the critical region for the same parameters as the black curve except the variances of error for all of the ToAs were changed from $(100\text{ns})^2$ to $(500\text{ns})^2$. It is clear that the dashed red curve nearly coincides with the black curve, indicating the variances of the error on the ToAs were greatly reduced by the averaging and did not affect the size

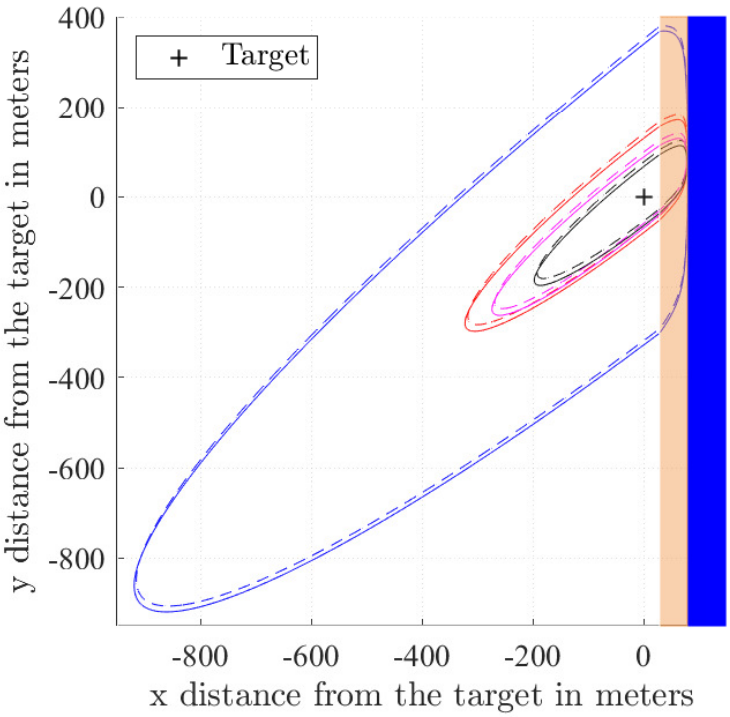

Figure 4.11: Critical regions with probability 0.99 for 1, 16, 49 and 30,000 transmissions: $\left(\sigma_{t_\ell}^{(n)}\right)^2 = \sigma_{t_\ell}^2 = (100 \text{ ns})^2$ and $\sigma_{x_\ell}^2 = \sigma_{y_\ell}^2 = (10 \text{ m})^2$.

of the critical region. This shows that for large N the size of the 0.99 probability critical region depends almost entirely on the variances of the measurement error on locations of the gateways, at least for the particular system used for Figs. 4.11, 4.12 and 4.13.

4.6 Extension to 3D Localization

Having completed the 2D analysis, it is straightforward to extend (3.33) to apply to a 3-dimensional space where the gateways and target have 3 co-ordinates. Let z be the third co-ordinate of the target's location and z_ℓ, $\ell = 1, 2, \ldots, L$, be the third co-ordinates of the gateways' locations. Then the 3-dimensional *a posteriori* density function is given as

$$f_{x,y,z\left|\widehat{\text{ToA}},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y\right.}(x,y,z\big|\boldsymbol{\tau},\widehat{\boldsymbol{g}}_x,\widehat{\boldsymbol{g}}_y) = \frac{\mathrm{e}^{-\frac{\gamma}{2}} f_{x,y,z}(x,y,z)}{\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} \mathrm{e}^{-\frac{\gamma}{2}} f_{x,y,z}(x,y,z)\mathrm{d}x\mathrm{d}y\mathrm{d}z}, \tag{4.5}$$

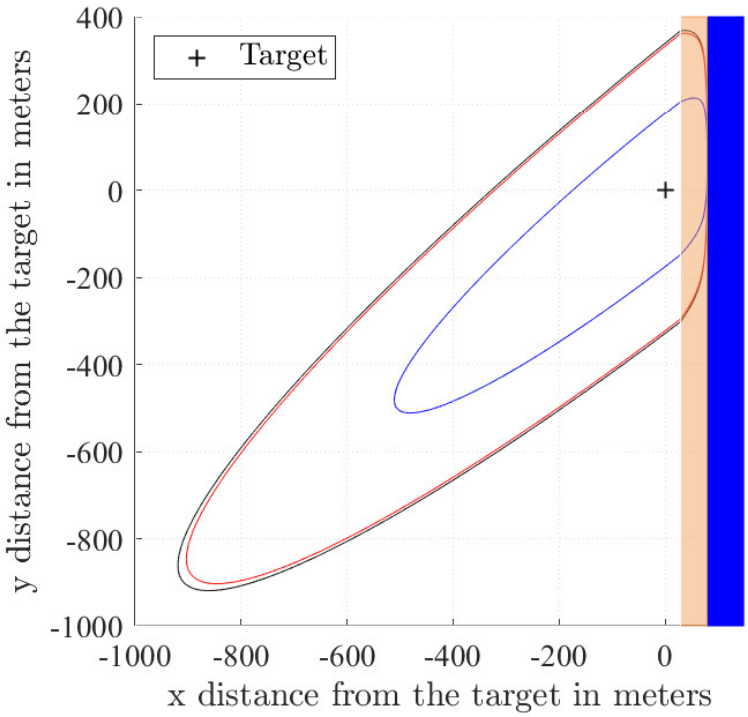

Figure 4.12: Critical regions with probability 0.99 for 1 transmission with different configurations.

where $f_{x,y,z}(x, y, z)$ is the *a priori* density of the target's co-ordinates and, for the 3-dimensional space, γ is now given by

$$\gamma = \sum_{\ell=1}^{L} \frac{\left(\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2 + (\widehat{z}_\ell - z)^2}/c\right)^2}{\sigma^2_{\mathsf{sum}_\ell}} - \sigma^2 \left(\sum_{\ell=1}^{L} \frac{\tau_\ell - \sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2 + (\widehat{z}_\ell - z)^2}/c}{\sigma^2_{\mathsf{sum}_\ell}} \right)^2, \tag{4.6}$$

where, for the 3-dimensional equation, $\sigma^2_{\mathsf{sum}_\ell}$ is given by (3.37).

It is very important to point out that, while (4.5) provides a 3-dimensional *a posteriori* pdf, it is capable of providing a 2-dimensional *a posteriori* pdf that is more compact, i.e., better, than (3.33). Equ. (4.5) is reduced to two dimensions by setting z to be the function of x and y that describes the elevation of a surface. Then (4.5) generates a 2-dimensional *a*

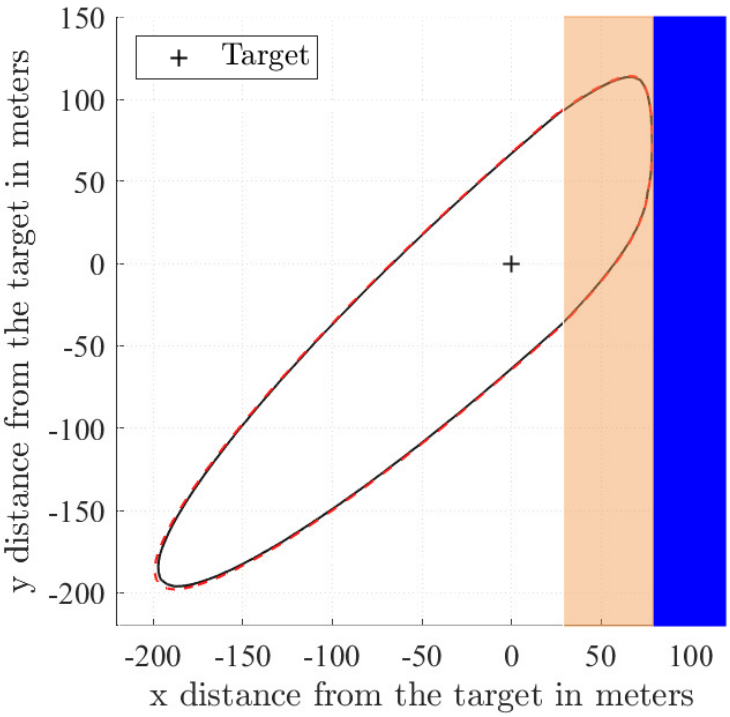

Figure 4.13: Critical regions with probability 0.99 for 30,000 transmission with different configurations.

posteriori pdf for a target that is known to sit on that surface. In fact, Equ. (4.5) can be reduced to (3.33) by setting z as well the vertical coordinates of all gateways to 0. However, leaving the vertical coordinates of the gateways as their true heights will generate a more accurate 2-dimensional *a posteriori* pdf than (3.33). The reason is that by assuming all the gateways are in the same horizontal plane introduces a bias in the time-of-flight equations and thereby into the estimates.

The worth of (4.5) is demonstrated with an example similar to the one that produced Fig. 4.4. The horizontal coordinates of the gateways are the same as in the 2-dimensional (2D) example, i.e., at $(\pm 100\,\mathrm{m}, \pm 100\,\mathrm{m})$, but now gateways 1, 2, 3, and 4 have heights 25 m, 85 m, 105 m and 0 m, respectively, with the gateway numbering following that of Fig. 4.1. Here, only the lucky observation from the ToA distributions with $\sigma_{t_\ell} = 25\mathrm{ns}$. The target is positioned at the origin, which is not where it was positioned for the 2D example, and a lower limit is placed on its vertical coordinate through an *a priori* pdf that is 0 for $z < -300\,\mathrm{m}$ and otherwise uniform.

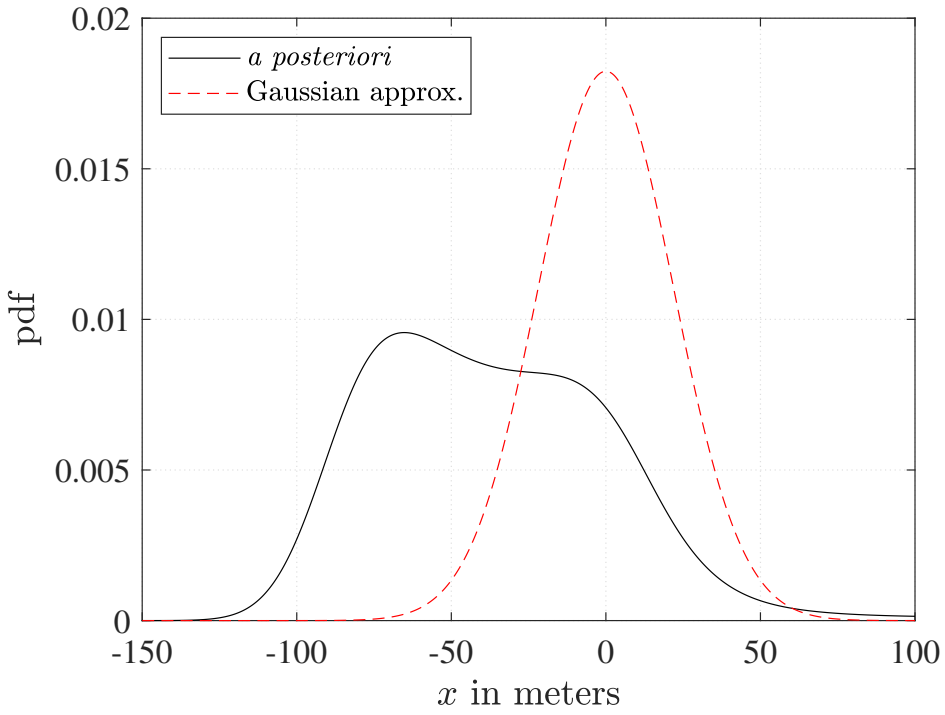

Figure 4.14: Marginal *a posteriori* pdf for the x coordinate computed from (4.5) and its corresponding Gaussian approximation for the lucky observation from an optimistic distribution for ToA error with $\sigma_{t_\ell} = 25\,\text{ns}$.

The three marginal *a posteriori* pdfs for the x, y and z coordinates along with their Gaussian approximations are plotted in Figs. 4.14, 4.15 and 4.16, respectively. Note that Gaussian approximations cannot accommodate an *a priori* pdf so they are not affected by it.

The three marginal *a posteriori* pdfs show that the ToF equations are certainly non-linear over the reach of the errors in the "optimistic" distribution, which has $\sigma_{t_\ell} = 25\text{ns}$, especially in the x and z directions. It is clear that the Foy's linearized model used to produce the Gaussian pdfs does not accurately represent the error in localization. When the standard deviation on the ToA errors is reduced to 10 ns, the linear model becomes reasonably accurate. All three of the marginal *a posteriori* pdfs are still noticeably non-symmetrical, but track the Gaussian approximations reasonably well.

Finally, the worth of using the (4.5) with a fixed value for the z coordinate instead of the 2D *a posteriori* pdf given by (3.33) is demonstrated. The pdfs were generated using (4.5) with z set to 0 m and (3.33) for the same system that produced Figs. 4.14, 4.15 and 4.16. Of

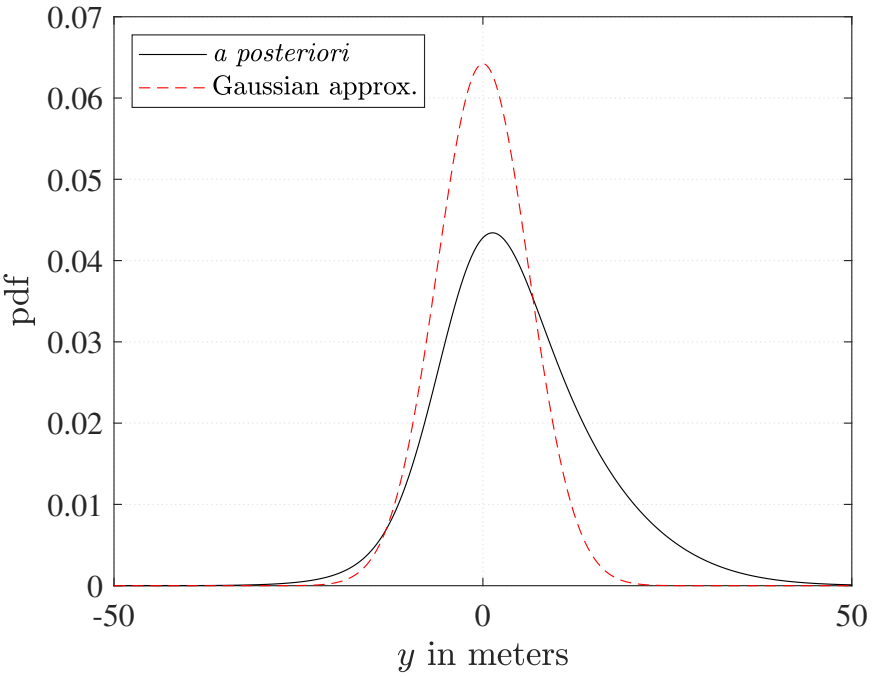

Figure 4.15: Marginal *a posteriori* pdf for the y coordinate computed from (4.5) and its corresponding Gaussian approximation for the lucky observation from an optimistic distribution for ToA error with $\sigma_{t_\ell} = 25\,\text{ns}$.

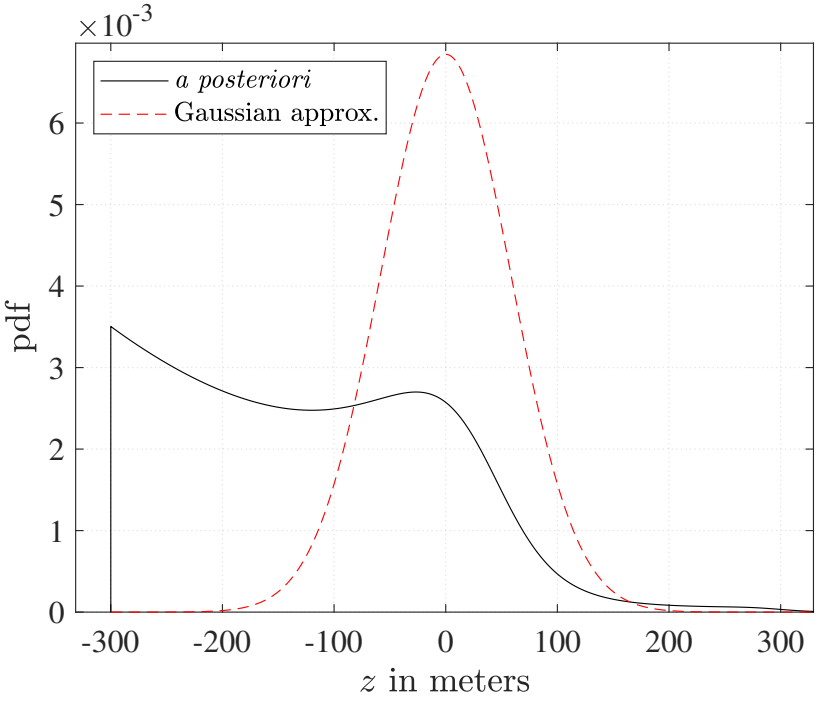

Figure 4.16: Marginal *a posteriori* pdf for the z coordinate computed from (4.5) and its corresponding Gaussian approximation for the lucky observation from an optimistic distribution for ToA error with $\sigma_{t_\ell} = 25\,\text{ns}$.

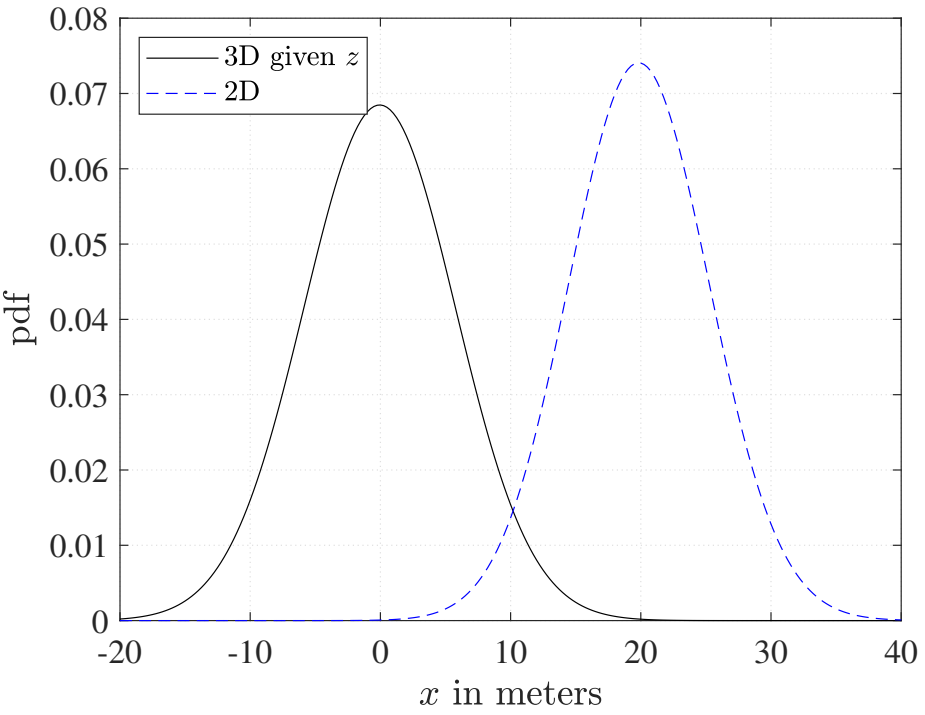

Figure 4.17: Marginal *a posteriori* pdfs for the x coordinate computed from (4.5) with z given to be $z = 0\,\text{m}$ and from (3.33) for the lucky observation from an optimistic distribution for ToA error with $\sigma_{\text{ToA}} = 25\,\text{ns}$.

course, the elevations of the gateways and target are ignored in (3.33). The resulting pdfs for the x-coordinate of the target are plotted in Fig. 4.17. There are two major differences between Figs. 4.14 and 4.17. First, (3.33) introduces a bias caused by the actual ToFs being larger than what were modeled. In this case the bias is 20 meters. Second, (3.33) is more sensitive to the target's horizontal position and produces a standard deviation that is smaller than it should be. It is easily concluded that (4.5) should always be used over (3.33).

5. *A Posteriori* PDF for Time-Difference of Arrival Measurements

5.1 Background

In the literature, localization systems based on time measurements of a signal transmitted from a source, i.e., target, and received at a set of spatially-separated receivers having known positions (or known with some measurement errors) [56] can be classified into two different groups distinguished by the types of signal transmitted by the source.

1. The first group centers on applications, like underwater sonar, where only statistical information on the radiating signal, e.g. its power spectral density, is available. For such localization systems, ToA measurement is not possible. Instead, TDoA measurements are made by pairwise cross-correlation of the signals received at L receivers [57]. A wealth of literature on TDoA measurements with pairwise cross-correlation can be found in [58].

2. In contrast, localization systems in modern wireless applications, such as Internet-of-Things (IoT), cannot measure TDoAs because the receivers (also known as gateways) in modern IoT systems (such as LoRaWAN) are not designed to relay the analog signals they receive to the application server for computing pairwise cross-correlation. Instead, the ToAs must be measured using a known portion of the transmitted signal, for example an embedded preamble. If used in a localization algorithm, the TDoA values are then computed from these ToA measurements.

It is pointed out that the statistics of the errors on the TDoA values calculated from ToA measurements will be different from TDoA measurements made by pairwise cross-

correlation. Therefore, literature that discusses a set of $(L-1)$ pairwise cross-correlation measurements does not directly apply to a set of $(L-1)$ TDoA values obtained by subtracting a reference ToA. For example, the work in [59] does not apply to applications considered in this paper. The authors in [59] show that, for the purpose of ordinary least squares (OLS) estimation, under special conditions for the signal and noise, the full set of $L(L-1)/2$ TDoA measurements obtained by pairwise cross-correlation can be transformed to a set of $(L-1)$ TDoA equations without loss of information.

This chapter has been included to support localization systems that fall into the first group, but the analysis is based on systems that fall into the second group, i.e., the analysis is based on ToA measurements made at the gateways. Within this second group of systems, there are two very different scenarios:

Scenario 1: the time that the message is transmitted from the target is measured and

Scenario 2: the time that the message is transmitted from the target is unknown.

Systems where the time of transmission is measured at the target, i.e., group 2 systems with scenario 1 transmission, is less common for two reasons: scenario 2 systems are much cheaper and there is very little difference in complexity/cost between measuring the time of transmission from a GPS signal and measuring the location from a GPS signal. Also, in both cases either the transmission time or the target's location has to be included in the message transmitted to the gateways.

Scenario 2 is the scenario of interest in this thesis. However, since there is one less independent measurement in scenario 2, the *a posteriori* pdf for scenario 2 will be more spread out and the marginal pdfs will have larger variances.

It is common to begin the analysis of systems with scenario 2 signals by using one ToA measurement as a reference to obtain $(L-1)$ time-difference-of-arrival (TDoA) values with respect to the remaining $(L-1)$ ToA measurements, where L is the total number of gateways. This is done to eliminate the unknown transmit time from $(L-1)$ equations. Each TDoA value can be translated into the difference in distances from the target to the corresponding

two gateways. Thus each TDoA value defines a hyperbola (or hyperboloid for 3-dimensional localization) that the target is supposed to lie on. Then the final estimation of the target's location is done based on the intersections of $(L-1)$ hyperbolas (hyperboloids) [31]. In particular, the maximum likelihood (ML) estimator based on TDoAs can be found in [31,60].

Another approach for Scenario 2 is to also estimate the target's transmit time as a nuisance parameter jointly with the source location. The maximum likelihood (ML) estimator is obtained in [44, 61], whereas the existence question for the ML estimator is recently examined in [62]. Since finding the ML solution is difficult, various simplifications are made (hence sub-optimal solutions are obtained) in [44, 61, 63] that are more amendable to fast computation.

It was observed via simulation results in [61] that the ML estimator obtained by jointly estimating the time the message was transmitted and the target's coordinates yields the same localization performance as the ML estimator obtained from TDoA values. The authors also pointed out that their simulation results show no performance difference when selecting different ToA measurements as a reference. They show clearly, however, that for sub-optimal estimators, such as linear least squares (LLS) and semi-definite programming (SDP), joint estimation of the transmit time and target's location based on ToAs is better than estimation of the target's location based on TDoAs. Similar observations were made in [44, 63].

All the above discussion pertains to the point estimation paradigm. Another paradigm for localization is region estimation. Each paradigm has its advantages. Point estimation provides estimates of the coordinates of the target, whereas region estimation, more specifically Bayesian region estimation, calculates the probability the target is in a specified region. While point estimation has been researched extensively [31], region estimation in the context of TDoA based localization has not.

5.2 The *a Posteriori* Pdf for Scenario 2 TDoAs

The objective of this section is to find the *a posteriori* pdf for the coordinates of a target in terms of the TDoAs calculated from the ToAs at a system of L spatially separated gateways.

The set of $(L-1)$ TDoAs is obtained by subtracting the ToA at gateway k from the $(L-1)$ ToAs at the other gateways. Specifically the set of $(L-1)$ TDoA estimates is given by

$$\widehat{\mathsf{TDoA}}_{\ell,k}(\zeta) = \widehat{\mathsf{ToA}}_{\ell}(\boldsymbol{\zeta}) - \widehat{\mathsf{ToA}}_{k}(\zeta), \quad \ell = 1, 2, \ldots, k-1, k+1, \ldots, L, \tag{5.1}$$

where gateway k is used as the reference.

Let $\tau_{\ell,k} = \widehat{\mathsf{TDoA}}_{\ell,k}(\zeta)$, then the equation above can be expressed as

$$\tau_{\ell,k} = \tau_{\ell} - \tau_{k}, \quad \ell = 1, 2, \ldots, k-1, k+1, \ldots, L.$$

The *a posteriori* pdf is given by (4.5) can be expressed in terms of the TDoAs by expressing γ in terms of TDoAs. The logic for converting the equation for γ given by (4.6) to an expression in terms of TDoAs is quite straightforward. From physics, it is clear that the *a posteriori* pdf given by (4.5) does not depend on the time of transmission. Therefore, the time of transmission could be advanced or retarded by any value without affecting the pdf given by (4.5). Therefore, after measuring τ_{ℓ}, $\ell = 1, 2, \ldots, L$, any values, say b, could be subtracted from all τ_{ℓ}, and then all occurrences of τ_{ℓ} in (4.6) could be replaced with $\tau_{\ell} - b$ without changing the *a posteriori* pdf. Now let $b = \tau_{k}$, then all occurrences of τ_{ℓ} in (4.6) must be replaced with $\tau_{\ell} - \tau_{k}$, which will be symbolized $\tau_{\ell,k}$, and (4.6) becomes

$$\begin{aligned}\gamma = &\sum_{\ell=1}^{L} \frac{\left(\tau_{\ell,k} - \sqrt{(\widehat{x}_{\ell} - x)^2 + (\widehat{y}_{\ell} - y)^2 + (\widehat{z}_{\ell} - z)^2}/c\right)^2}{\sigma^2_{\mathsf{sum}_{\ell}}} \\ &- \sigma^2 \left(\sum_{\ell=1}^{L} \frac{\tau_{\ell,k} - \sqrt{(\widehat{x}_{\ell} - x)^2 + (\widehat{y}_{\ell} - y)^2 + (\widehat{z}_{\ell} - z)^2}/c}{\sigma^2_{\mathsf{sum}_{\ell}}} \right)^2 .\end{aligned} \tag{5.2}$$

Note that $\tau_{\ell,k}$ for $\ell = k$ is 0 so γ is a function of $\tau_{\ell,k}$, $\ell = 1, 2, \ldots, k-1, k+1, \ldots, L$.

Therefore, the *a posteriori* pdf given by (4.5) can be computed using TDoAs $\tau_{\ell,k}$ by computing γ with (5.2). This result is also proved in Appendix C (page 105) in a different way.

5.3 Applicability of the *a Posteriori* Pdf in Applications where ToAs Cannot Be Estimated

There are applications where it is not possible to estimate the ToAs. Such applications include underwater sonar where the target emits a continuous random sonar signal. In these applications it is possible to estimate TDoAs using the peak value of a cross-correlation [57]. Typically the continuous signal is windowed in time at all of the receivers and then these finite time signals are cross-correlated and the peak of the cross-correlation becomes the TDoA estimate.

Since the parameters in (5.2) are the variances of the errors in the ToA estimates, these variances must be known to find the *a posteriori* with TDoA estimates. This begs the question "If the ToAs cannot be estimated, how can the variances of the errors on the ToA estimates be found?". It turns out, unlike the ToA estimates themselves, the variances of the errors on the ToA estimates can be calculated from the power spectral density functions of the random signal and the random noises corrupting that signal at different receivers. The adaptability of (5.2) to applications where the signal emitted/transmitted by the target is continuous and random is explored below. It is pointed out that similar techniques and variance analysis of the TDoA estimates by cross-correlation, but based on samples (discrete-time) of the received signals can be found in [59, 64].

Suppose a system of spatially separated receivers, which have been referred to as gateways up to this point, receive a continuous random signal from a target and the receivers relay these signals to a central processor. Furthermore, suppose the propagation delays in the links from the receivers to the central processor are known and the auto-correlation function and power spectral density of the noise corrupting the inputs to the receivers are also known by the central processor. The central processor can then correct for the propagation delays incurred in the links from the receivers. Then the central processor is in a position to window the received signals and cross-correlate the windowed segments. Obviously, the location of the peaks in the cross-correlation are the TDoA estimates.

Let the signal emitted/transmitted by the target be sample function $s(t, \zeta)$ from stochas-

tic process $s(t, \boldsymbol{\zeta})$. That sample function is unknown to the receivers, but the auto-correlation function and power spectral density of stochastic process $s(t, \boldsymbol{\zeta})$ are known to the receivers. Let the signal plus noise received by gateway ℓ, i.e., receiver ℓ, be denoted as

$$v_\ell(t, \zeta) = \alpha_\ell s(t - d_\ell, \zeta) + n_\ell(t, \zeta), \tag{5.3}$$

where α_ℓ is the linear gain of the link between the target and gateway ℓ, d_ℓ is the unknown propagation between the target and gateway ℓ and $n_\ell(t, \zeta)$, which is a sample function from stochastic process $n_\ell(t, \boldsymbol{\zeta})$, is the noise corrupting the signal received by gateway ℓ. Furthermore $n_\ell(t, \boldsymbol{\zeta})$ is independent of $s(t - d_\ell, \boldsymbol{\zeta})$ and $n_k(t, \boldsymbol{\zeta})$ for $k \neq \ell$. The central processor windows the signal from the gateways with a window of length T, where T is sufficiently large for cross correlation to produce a good estimate.

Note, in order to keep the number of symbols used in this thesis as small as possible, the symbols used for the windowed signals will be the same as those used for the continuous signals.

The correlator has three inputs: two are time signals and the other controls the delay that the correlator forces on one of the time signal inputs. The auto correlation function for the stochastic process from which time signal $s(t, \zeta)$ is drawn is denoted $R_{ss}(t_1, t_2)$, where the auto-correlation function is defined as

$$R_{ss}(t_1, t_2) = \mathbb{E}[s(t_1, \boldsymbol{\zeta}), s^*(t_2, \boldsymbol{\zeta})].$$

Here $s(t_1, \boldsymbol{\zeta})$ and $s(t_2, \boldsymbol{\zeta})$ are random variables that take on the values of $s(t, \zeta)$ at times t_1 and t_2, respectively.

The output of the correlator when the inputs are from receivers ℓ and k is

$$C_{\ell,k}(\tau, \boldsymbol{\zeta}) = \int_{-\infty}^{\infty} \big(\alpha_k s(t - d_k, \boldsymbol{\zeta}) + n_k(t, \boldsymbol{\zeta})\big)\big(\alpha_\ell s(t - d_\ell - \tau, \boldsymbol{\zeta}) + n_\ell(\tau, \boldsymbol{\zeta})\big)\mathrm{d}t, \tag{5.4}$$

where the signals in the integrand of (5.4) are the windowed signals so the limits of the

integral could be finite and the delay of τ is imposed on the signal from the receiver ℓ.

The value of τ that maximizes $C_{\ell,k}(\tau, \boldsymbol{\zeta})$ is taken to be the estimate $\widehat{\mathsf{TDoA}}_{\ell,k}(\zeta)$, which in this case is

$$\tau_{@\mathsf{max}} = \widehat{\mathsf{TDoA}}_{\ell,k}(\zeta) = \tau_{\ell,k} = d_\ell - d_k + \Delta\tau_{\ell,k}(\zeta),$$

where $\tau_{\ell,k}$ is short for $\widehat{\mathsf{TDoA}}_{\ell,k}(\zeta)$ and $\Delta\tau_{\ell,k}(\zeta)$ is the error in the TDoA estimate for outcome ζ.

In this analysis we are only interested in the statistics of the error $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$. Without loss of generality, d_k and d_ℓ can be set to zero. This is done to remove symbolic clutter as well as remove a level of abstraction. This can be done in this analysis since $d_\ell - d_k$ is of no interest. Then $\tau_{@\mathsf{max}} = \Delta\tau_{\ell,k}(\boldsymbol{\zeta})$. Also, without loss of generality, the window is centered at the time origin and extends from $t = -T/2$ to $t = T/2$.

There are two ways to deal with windowed signals. One way is to assume that the portion outside the window is zero and the other is to assume the windowed signal is an excerpt from a periodic signal with period T. Neither assumption is perfect and both introduce error when a shifted version of the signal is the integrand of an integral over an interval of T. In this analysis the windowed signal will be converted to a continuous infinite length periodic signal by repeatedly pre-pending and appending the windowed segment to itself.

After setting $d_\ell = 0$ for all ℓ and extending the windowed signals to make them periodic, the cross correlation function given by (5.4) can be written as

$$C_{\ell,k}(\Delta\tau, \boldsymbol{\zeta}) = \mathsf{signal} + \mathsf{noise}_1 + \mathsf{noise}_2 + \mathsf{noise}_3, \tag{5.5}$$

where

$$\mathsf{signal} = \int_{-T/2}^{T/2} \alpha_k \alpha_\ell s(t,\zeta) s(t-\Delta\tau,\zeta)\mathrm{d}t,$$

$$\mathsf{noise}_1 = \int_{-T/2}^{T/2} \alpha_k s(t,\zeta) n_\ell(t-\Delta\tau,\zeta)\mathrm{d}t,$$

$$\mathsf{noise}_2 = \int_{-T/2}^{T/2} \alpha_\ell n_k(t,\zeta) s(t-\Delta\tau,\zeta)\mathrm{d}t,$$

$$\mathsf{noise}_3 = \int_{-T/2}^{T/2} n_k(t,\zeta) n_\ell(t-\Delta\tau,\zeta)\mathrm{d}t,$$

where the delay introduced by the correlator has been changed from τ to $\Delta\tau$. Since all d_ℓ have been set to zero, $\Delta\tau_{@\mathsf{max}}$ is the estimation error, i.e., $\Delta\tau_{\ell,k}(\zeta) = \Delta\tau_{@\mathsf{max}}$.

The windowed time signals have been extended to be periodic with period T so they can be expressed as their Fourier Series, which is

$$s(t,\zeta) = \sum_{m=0}^{M} A_s(m,\zeta)\cos\left(\frac{2\pi m t}{T} + \phi_s(m,\zeta)\right)$$

and

$$n_\ell(t,\zeta) = \sum_{m=0}^{M} A_\ell(m,\zeta)\cos\left(\frac{2\pi m t}{T} + \phi_\ell(m,\zeta)\right)$$

where $A_s(m,\zeta)$ and $\phi_s(m,\zeta)$ are the Fourier series coefficients for amplitude and phase for the sample function $s(t,\zeta)$ and $A_\ell(m,\zeta)$ and $\phi_\ell(m,\zeta)$ are the Fourier Series coefficients for amplitude and phase for the sample function $n_\ell(t,\zeta)$.

Then replacing $s(t,\zeta)$ and $n_k(t,\zeta)$ and $n_\ell(t,\zeta)$ in (5.5) with their Fourier series, changing

the order of summation and integration and integrating produces:

$$\begin{aligned}
\mathsf{signal} &= \alpha_\ell \alpha_k \sum_{m=0}^{M} \frac{A_s^2(m,\zeta)}{2} T \cos\Big(\frac{2\pi m}{T}\Delta\tau\Big),\\
\mathsf{noise}_1 &= \alpha_k \sum_{m=0}^{M} \frac{A_s(m,\zeta)A_\ell(m,\zeta)}{2} \cos\Big(\frac{2\pi m}{T}\Delta\tau + \phi_\ell(m,\zeta) - \phi_s(m,\zeta)\Big),\\
\mathsf{noise}_2 &= \alpha_\ell \sum_{m=0}^{M} \frac{A_s(m,\zeta)A_k(m,\zeta)}{2} \cos\Big(\frac{2\pi m}{T}\Delta\tau + \phi_s(m,\zeta) - \phi_k(m,\zeta)\Big),\\
\mathsf{noise}_3 &= \sum_{m=0}^{M} \frac{A_\ell(m,\zeta)A_k(m,\zeta)}{2} \cos\Big(\frac{2\pi m}{T}\Delta\tau + \phi_\ell(m,\zeta) - \phi_k(m,\zeta)\Big).
\end{aligned}$$

It is not possible to estimate the ToAs in practice since $s(t,\zeta)$ is drawn from a stochastic process and therefore, unknown to the central processor. However, in a laboratory setting $s(t,\zeta)$ can be fed into the reference input of the correlator. In this case the $\Delta\tau$ that maximizes the correlator output, i.e., $\Delta\tau_{@\mathsf{max}}$, is the error on the estimated ToA for receiver ℓ, which is denoted $\Delta\tau_\ell(\zeta)$. It is worth repeating that since all d_ℓ have been set to zero, $\Delta\tau_{@\mathsf{max}}$ is the estimation error, i.e., $\Delta\tau_\ell(\zeta) = \Delta\tau_{@\mathsf{max}}$.

Setting $\alpha_k = 1$ and $n_k(t,\zeta) = 0$ in (5.5) changes $\Delta\tau_{@\mathsf{max}}$ from being $\Delta\tau_{\ell,k}(\zeta)$ to being $\Delta\tau_\ell(\zeta)$. Therefore, (5.5) can be used to find the error in estimates of ToAs or TDoAs, at least in a laboratory setting.

5.3.1 Variances of ToAs and TDoAs

The variances of $\Delta\tau_\ell(\boldsymbol{\zeta})$ and $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$ can be found approximately through the following sequence of steps.

Step 1: Approximate the signal in (5.5) with its second order Maclaurin polynomial in $\Delta\tau$ (i.e., the second order Taylor series about $\Delta\tau = 0$). For "signal" the correlator output will be maximum at $\Delta\tau = 0$. Therefore, $\frac{\mathrm{d}\,\mathsf{signal}}{\mathrm{d}\Delta\tau} = 0$ at $\Delta\tau = 0$ and the Maclaurin

polynomial becomes

$$\mathsf{signal} \simeq \mathsf{signal}\Big|_{\Delta\tau=0} + \frac{\mathrm{d}^2\mathsf{signal}}{\mathrm{d}\Delta\tau^2}\Big|_{\Delta\tau=0} \frac{\Delta\tau^2}{2},$$

where signal is the output of the correlator as a function $\Delta\tau$ when the inputs are $\alpha_k s(t,\zeta)$ and $\alpha_\ell s(t,\zeta)$. Since the output of the correlator is a function of $\Delta\tau$ and ζ and the two inputs are scaled copies of sample function $s(t,\zeta)$, it will be denoted $C_{ss}(\Delta\tau,\zeta)$. Using this notation the approximation above is written:

$$\mathsf{signal} \simeq C_{ss}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0} + C_{ss}^{"}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0} \frac{\Delta\tau^2}{2}. \tag{5.6}$$

Note that $C_{ss}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0}$ is the energy in $\alpha_k\alpha_\ell s(t,\zeta)$.

Step 2: Set up the equation that, when solved, produces $\Delta\tau_{\ell,k}(\zeta)$. This is done by replacing signal in (5.5) with the approximation given by (5.6) then taking the derivative of (5.5) with respect to $\Delta\tau$ and setting it equal to zero. After moving the signal component to the left side the result is

$$C_{ss}^{"}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0} \Delta\tau_{\ell,k}(\zeta) = \left(-\frac{\mathrm{d}\ \mathsf{noise}_1}{\mathrm{d}\Delta\tau} - \frac{\mathrm{d}\ \mathsf{noise}_2}{\mathrm{d}\Delta\tau} - \frac{\mathrm{d}\ \mathsf{noise}_3}{\mathrm{d}\Delta\tau} \right)\Bigg|_{\Delta\tau=\Delta\tau_{\ell,k}(\boldsymbol{\zeta})}, \tag{5.7}$$

where

$$\frac{\mathrm{d}\ \mathsf{noise}_1}{\mathrm{d}\Delta\tau} = -\alpha_k \sum_{m=0}^{M} \frac{A_s(m,\zeta)A_\ell(m,\zeta)}{2}\frac{2\pi m}{T}\times \sin\left(\frac{2\pi m}{T}\Delta\tau + \phi_\ell(m,\zeta) - \phi_s(m,\zeta)\right),$$

$$\frac{\mathrm{d}\ \mathsf{noise}_2}{\mathrm{d}\Delta\tau} = -\alpha_\ell \sum_{m=0}^{M} \frac{A_s(m,\zeta)A_k(m,\zeta)}{2}\frac{2\pi m}{T}\times \sin\left(\frac{2\pi m}{T}\Delta\tau + \phi_s(m,\zeta) - \phi_k(m,\zeta)\right),$$

$$\frac{\mathrm{d}\ \mathsf{noise}_3}{\mathrm{d}\Delta\tau} = -\sum_{m=0}^{M} \frac{A_\ell(m,\zeta)A_k(m,\zeta)}{2}\frac{2\pi m}{T}\times \sin\left(\frac{2\pi m}{T}\Delta\tau + \phi_\ell(m,\zeta) - \phi_k(m,\zeta)\right),$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\Delta\tau^2}C_{ss}(\Delta\tau,\zeta) = -\alpha_k\alpha_\ell \sum_{m=0}^{M} \frac{A_s^2(m,\zeta)}{2}\left(\frac{2\pi m}{T}\right)^2 \cos\left(\frac{2\pi m}{T}\Delta\tau\right) \text{ and },$$

$$\frac{\mathrm{d}^2}{\mathrm{d}\Delta\tau^2}C_{ss}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0} = -\alpha_k\alpha_\ell \sum_{m=0}^{M} \frac{A_s^2(m,\zeta)}{2}\left(\frac{2\pi m}{T}\right)^2. \tag{5.8}$$

The equation above plays a critical role in the quest to find the variances of $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$ and $\Delta\tau_\ell(\boldsymbol{\zeta})$, but the specific values of $\Delta\tau_{\ell,k}(\zeta)$ are not needed. Therefore, the equation above does not need to be solved.

Step 3: Find the mean of $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$. Since $C_{ss}(\Delta\tau)$ is an even function of $\Delta\tau$, the pdf of $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$ will be an even function and the mean of $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$ will be zero.

Step 4: Find $\mathbb{E}[\Delta\tau_{\ell,k}^2(\boldsymbol{\zeta})]$. This expectation is the variance of $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$ since the mean of $\Delta\tau_{\ell,k}(\boldsymbol{\zeta})$ is zero. This is done by treating the sample functions in (5.7) as stochastic processes, squaring (5.7) and taking the expectation of both sides.

First, find the expectation of the left-hand side. This is done by assuming the signal is somewhat ergodic in the sense that the amplitude coefficients, $A_s(m,\zeta)$ are very weakly dependent on ζ (however, the phase coefficients could depend on ζ). In which case, i.e., evaluate this at $\Delta\tau = 0$, both $C_{ss}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0}$ and $C_{ss}^{"}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0}$ weakly

depend on ζ and can be treated as constants. Under this assumption,

$$C_{ss}^{"}(\Delta\tau,\zeta)\Big|_{\Delta\tau=0} = -\alpha_k\alpha_\ell \sum_{m=0}^{M} \frac{A_s^2(m)}{2}(2\pi m/T)^2 \quad \forall\zeta,$$

where $A_s(m,\zeta)$ is assumed to be independent of ζ and therefore written $A_s(m)$.

Therefore, the left-side of the square of (5.7) is

$$\big(C_{ss}^{"}(\Delta\tau,\boldsymbol{\zeta})\big)^2\Big|_{\Delta\tau=0} \Delta\tau_{\ell,k}^2(\boldsymbol{\zeta}) = \left(-\alpha_k\alpha_\ell \sum_{m=0}^{M} \frac{A_s^2(m)}{2}(2\pi m/T)^2 \right)^2 \Delta\tau_{\ell,k}^2(\boldsymbol{\zeta}). \tag{5.9}$$

The square of the right side of (5.7) has six different terms: three of which are the squares of the 3 terms on the right-side of (5.7) and 3 of which are the products of different terms. The expectation of a product of functions can be distributed to the functions of independent random variables. Doing this has an expectation of a function whose mean is zero. One of the expectations being zero ensures the product is zero and therefore ensures the three terms with "cross products" will be zero.

This leaves the 3 squared terms. The expectations of which are

$$\mathbb{E}\left[\left(\frac{\mathsf{dnoise}_1}{\mathrm{d}\Delta\tau}\right)^2\right] = \tfrac{1}{2}\alpha_k^2 \textstyle\sum_{m=0}^{M} \mathbb{E}\left[\frac{A_s^2(m,\zeta)}{2}\right] \mathbb{E}\left[\frac{A_\ell^2(m,\zeta)}{2}\right]\left(\tfrac{2\pi}{T}m\right)^2, \tag{5.10}$$

$$\mathbb{E}\left[\left(\frac{\mathsf{dnoise}_2}{\mathrm{d}\Delta\tau}\right)^2\right] = \tfrac{1}{2}\alpha_\ell^2 \textstyle\sum_{m=0}^{M} \mathbb{E}\left[\frac{A_s^2(m,\zeta)}{2}\right] \mathbb{E}\left[\frac{A_k^2(m,\zeta)}{2}\right]\left(\tfrac{2\pi}{T}m\right)^2, \tag{5.11}$$

$$\mathbb{E}\left[\left(\frac{\mathsf{dnoise}_3}{\mathrm{d}\Delta\tau}\right)^2\right] = \tfrac{1}{2} \textstyle\sum_{m=0}^{M} \mathbb{E}\left[\frac{A_\ell^2(m,\zeta)}{2}\right] \mathbb{E}\left[\frac{A_k^2(m,\zeta)}{2}\right]\left(\tfrac{2\pi}{T}m\right)^2. \tag{5.12}$$

Suppose the bandwidth of the signal, which is $\frac{2\pi M}{T}\,\frac{\mathsf{rad}}{\mathsf{sec}}$, is much larger than $\frac{2\pi}{T}$. That is suppose the time-bandwidth product is much much greater than 2π, i.e., $T\times\frac{2\pi M}{T} \gg 2\pi$ or $M \gg 1$. Then the frequency spacing of the Fourier sinusoids, which is $\frac{2\pi}{T}\,\frac{\mathsf{rad}}{\mathsf{sec}}$, is small and the power spectral densities of $s(t,\boldsymbol{\zeta})$ and $n_\ell(t,\boldsymbol{\zeta})$ are approximately

$$\mathcal{S}_{ss}\left(\frac{2\pi m}{T}\right) \simeq \frac{\mathbb{E}\left[\frac{A_s^2(m,\zeta)}{2}\right]}{\frac{2\pi}{T}} \text{ and } \mathcal{S}_{n_\ell n_\ell}\left(\frac{2\pi m}{T}\right) \simeq \frac{\mathbb{E}\left[\frac{A_\ell^2(m,\zeta)}{2}\right]}{\frac{2\pi}{T}},$$

which means

$$\mathbb{E}\left[\frac{A_s^2(m,\zeta)}{2}\right] \simeq \frac{2\pi}{T}\mathcal{S}_{ss}\left(\frac{2\pi m}{T}\right) \text{ and } \mathbb{E}\left[\frac{A_\ell^2(m,\zeta)}{2}\right] \simeq \frac{2\pi}{T}\mathcal{S}_{n_\ell n_\ell}\left(\frac{2\pi m}{T}\right).$$

Then the summations in (5.10), (5.11) and (5.12) can be treated as Riemann sums and these equations become

$$\mathbb{E}\left[\left(\frac{\mathrm{d}\ \mathsf{noise}_1}{\mathrm{d}\Delta\tau}\right)^2\right] = \tfrac{1}{2}\alpha_k^2\left(\tfrac{T}{2\pi}\right)^3 \int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathcal{S}_{n_\ell n_\ell}(\omega)\mathrm{d}\,\omega, \tag{5.13}$$

$$\mathbb{E}\left[\left(\frac{\mathrm{d}\ \mathsf{noise}_2}{\mathrm{d}\Delta\tau}\right)^2\right] = \tfrac{1}{2}\alpha_\ell^2\left(\tfrac{T}{2\pi}\right)^3 \int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathcal{S}_{n_k n_k}(\omega)\mathrm{d}\omega, \tag{5.14}$$

$$\mathbb{E}\left[\left(\frac{\mathrm{d}\ \mathsf{noise}_3}{\mathrm{d}\Delta\tau}\right)^2\right] = \tfrac{1}{2}\left(\tfrac{T}{2\pi}\right)^3 \int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{n_\ell n_\ell}(\omega)\mathcal{S}_{n_k n_k}(\omega)\mathrm{d}\omega. \tag{5.15}$$

For the same reason $\mathbb{E}[C"(\Delta\tau,\zeta)]\Big|_{\Delta\tau=0}$ can be expressed as

$$\left(\alpha_k\alpha_\ell\left(\frac{T}{2\pi}\right)^2 \int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathrm{d}\omega\right)^2.$$

Then, after squaring (5.7), taking the expectation of both sides, and dividing both sides by $\big(C"(\Delta\tau,\zeta)\big)^2$, has

$$\begin{aligned}\mathbb{E}[\tau_{\ell,k}^2(\zeta)] = \frac{\pi}{T}\frac{\left[\alpha_\ell^2 \int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathcal{S}_{n_\ell n_\ell}(\omega)\mathrm{d}\omega + \alpha_k^2 \int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathcal{S}_{n_k n_k}(\omega)\mathrm{d}\omega\right]}{\alpha_\ell^2\alpha_k^2\left(\int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathrm{d}\omega\right)^2} \\ + \frac{\pi}{T}\frac{\int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{n_\ell n_\ell}(\omega)\mathcal{S}_{n_k n_k}(\omega)\mathrm{d}\omega}{\alpha_\ell^2\alpha_k^2\left(\int_{\omega=0}^{2\pi M/T} \omega^2 \mathcal{S}_{ss}(\omega)\mathrm{d}\omega\right)^2}.\end{aligned} \tag{5.16}$$

For example, if $\mathcal{S}_{ss}(\omega)$, $\mathcal{S}_{n_k n_k}(\omega)$ and $\mathcal{S}_{n_\ell n_\ell}(\omega)$ are all flat, low pass power spectral densities with spectral constants 1, β_k and β_ℓ, respectively, all with unit watts/Hz

and all stochastic processes have the same bandwidth B Hz (or $2\pi B$ rad/sec) then $\mathbb{E}[\tau^2_{\ell,k}(\boldsymbol{\zeta})]$ is

$$\begin{aligned}\mathbb{E}[\tau^2_{\ell,k}(\boldsymbol{\zeta})] &= \frac{\pi}{T}\frac{\int_0^{2\pi B}\omega^2(\alpha_\ell^2\beta_k+\alpha_k^2\beta_\ell+\beta_\ell\beta_k)\mathrm{d}\omega}{\alpha_\ell^2\alpha_k^2\Big(\int_0^{2\pi B}\omega^2\mathrm{d}\omega\Big)^2}\\ &= \frac{\pi}{T}\frac{(\alpha_\ell^2\beta_k+\alpha_k^2\beta_\ell+\beta_\ell\beta_k)}{\alpha_\ell^2\alpha_k^2\frac{1}{3}(2\pi B)^3}\\ &= \frac{3}{8\pi^2B^3T}\left(\frac{\beta_k}{\alpha_k^2}+\frac{\beta_\ell}{\alpha_\ell^2}+\frac{\beta_k\beta_\ell}{\alpha_\ell^2\alpha_k^2}\right)\\ &= \frac{3}{8\pi^2B^3T}\left(\frac{1}{\mathsf{SNR}_k}+\frac{1}{\mathsf{SNR}_\ell}+\frac{1}{\mathsf{SNR}_k\mathsf{SNR}_\ell}\right),\end{aligned} \tag{5.17}$$

where SNR_k and SNR_ℓ are the ratios of signal power to noise power at receivers k and ℓ, respectively.

Step 5: Find the variance for the time of arrival at receiver ℓ. This is done by setting $n_k(t,\zeta)$ to zero of all ζ and setting α_k to 1 and using (5.16). Since $n_k(t,\zeta)=0\ \forall\zeta$, $\mathcal{S}_{n_kn_k}(\omega)=0$. Substituting 0 for $\mathcal{S}_{n_kn_k}(\omega)$ and 1 for α_k in (5.16) produces the variance for ToA at receiver ℓ. It is given by

$$\mathbb{E}[\tau^2_\ell(\boldsymbol{\zeta})] = \frac{\pi}{T}\frac{\int_0^{2\pi B}\omega^2\mathcal{S}_{ss}(\omega)\mathcal{S}_{n_\ell n_\ell}(\omega)\mathrm{d}\omega}{\big(\alpha_\ell^2\int_0^{2\pi B}\omega^2\mathcal{S}_{ss}(\omega)\mathrm{d}\omega\big)^2}. \tag{5.18}$$

Under the same assumptions with the above example for calculating $\mathbb{E}[\tau^2_{\ell,k}(\boldsymbol{\zeta})]$, $\mathbb{E}[\tau^2_\ell(\boldsymbol{\zeta})]$ is

$$\mathbb{E}[\tau^2_{\ell,k}(\boldsymbol{\zeta})] = \frac{\pi}{T}\frac{\int_0^{2\pi B}\omega^2\beta_\ell\mathrm{d}\omega}{\alpha_\ell^2\left(\int_0^{2\pi B}\omega^2\mathrm{d}\omega\right)^2} = \frac{\pi\beta_\ell}{T\alpha_\ell^2\frac{1}{3}(2\pi B)^3} = \frac{3}{8\pi^2TB^3\mathsf{SNR}_\ell}, \tag{5.19}$$

where SNR_k and SNR_ℓ are the ratios of signal power to noise power at receivers k and ℓ, respectively.

5.4 Verification of Equations for $\mathbb{E}[\tau_{\ell,k}^2(\boldsymbol{\zeta})]$ and $\mathbb{E}[\tau_{\ell}^2(\boldsymbol{\zeta})]$

In this section the equation for $\mathbb{E}[\tau_{\ell,k}^2(\boldsymbol{\zeta})]$ and $\mathbb{E}[\tau_{\ell}^2(\boldsymbol{\zeta})]$, i.e., (5.16) and (5.18), are verified for a random signal and random noise with the following power spectral densities:

$$\mathcal{S}_{ss}(\omega) = \begin{cases} 1\ \frac{W}{Hz}, & 0 \leq \omega \leq 2\pi B \\ 0, & \text{otherwise,} \end{cases},$$
$$\mathcal{S}_{n_\ell n_\ell}(\omega) = \frac{\mathcal{S}_{ss}(\omega)}{\mathsf{SNR}} \text{ for all } \ell,$$

where $B = 400$Hz and SNR is a parameter of the experiment.

The cross correlator is simulated and $\tau_{\ell,k}(\zeta)$ and $\tau_{\ell}(\zeta)$ are found for 10,000 outcomes. The means and variances are computed from these outcomes and compared to the theoretical results.

The parameters used in verification are chosen to test the sensitivities of the assumptions made in the derivation of the variance equations. The two assumptions were T is chosen to be much greater than $\frac{1}{B}$ and the SNR is sufficiently high for $\Delta\tau_{\ell,k}(\zeta)$ to be small enough for a second order Maclaurin polynomial to be used for $C_{\ell,k}\big(\Delta\tau_{\ell,k}(\zeta), \zeta\big)$. The results are tabulated in Tables 5.1 and 5.2.

Comparison of the standard deviations calculated from the theoretically derived (5.17) and (5.19) to the standard deviations obtained experimentally show a reasonably good agreement. In most cases the agreement is within 10%. The case where the SNR $= -10$ dB and $T \times B = 10$, which stress the assumption of high time-bandwidth product and reasonable SNR. In that case the theoretical standard deviation differs by about 15%.

While limited in number, the experimental results captured in Tables 5.1 and 5.2 indicate Equations (5.17) and (5.19) are reasonable and can be used to draw conclusions on the differences between TDoA valued obtained directly through cross-correlation and those obtained by taking the difference of ToAs.

Table 5.1: An example for the estimates of the standard deviation for TDoA in units seconds. The experimental standard deviations were calculated from $10,000$ cross-correlation measurements.

SNR		$T \times B = 10$	$T \times B = 100$
-10 dB	Theoretical	0.00062262	0.00022042
	Experimental	0.00067063	0.00019689
0 dB	Theoretical	0.00026691	0.00008440
	Experimental	0.00028146	0.00008393
10 dB	Theoretical	0.00013189	0.00004170
	Experimental	0.00012384	0.00004177

Table 5.2: An example for the estimates of the standard deviation for ToA in units seconds. The experimental standard deviations were calculated from $10,000$ cross-correlation measurements.

SNR		$T \times B = 10$	$T \times B = 100$
-10 dB	Theoretical	0.00027403	0.000086657
	Experimental	0.00031921	0.000087236
0 dB	Theoretical	0.00015410	0.000048731
	Experimental	0.00015150	0.000048136
10 dB	Theoretical	0.00008665	0.000027403
	Experimental	0.00008188	0.000027471

5.5 Applicability of (3.33) and (4.5) to Systems that Measure TDoAs and Not ToAs

In general (3.33) and (4.5) cannot be applied to systems that measure the TDoAs directly using cross-correlation. However, there are situations where (3.33) and (4.5) can be used to generate an applicable *a posteriori* pdf. The argument for these two assertions is provided as a sequence of assertions arising from the expressions for variance of error that were developed in this chapter.

The logical argument for how and when to use (3.33) and (4.5) to get an approximate *a posteriori* pdf for systems that estimate the TDoAs directly using cross-correlation of windowed segments of continuous random signals is provided as the sequence of assertions given below:

1. It is reasonably assumed that the power spectral density of the noise corrupting the received signal is known. In other words it is assumed the power spectral densities of both the Johnson noise for the receiver as well as the environmental ambient noise picked up by the antenna/sensor can be either measured or calculated.

2. It is reasonably assumed that the power spectral density of the continuous signal plus noise picked up by a receiver can be measured.

3. The power spectral density of the signal can be calculated by subtracting the power spectral density of the noise from the power spectral density measured in assertion 2.

4. Since the power spectral densities of the transmitted signal and noise are available, the variance of error for the cross-correlation based TDoA estimator can be calculated from (5.16). The variance depends on T, α_ℓ, α_k, $\mathcal{S}_{ss}(\omega)$, $\mathcal{S}_{n_k n_k}(\omega)$ and $\mathcal{S}_{n_\ell n_\ell}(\omega)$.

5. Even though $s(t,\zeta)$ is not known and the ToAs cannot be estimated, the variance of error on the ToAs estimated from a cross-correlation with $s(t,\zeta)$ can be found using (5.18). From (5.18) the variance depends on T, α_ℓ, $\mathcal{S}_{ss}(\omega)$ and $\mathcal{S}_{n_\ell n_\ell}(\omega)$.

6. Comparing (5.16) and (5.18) shows the variance of the error on the TDoA estimates is related to the variance of error on the ToA estimates (if it were possible to estimate

ToAs) by

$$\mathbb{E}[\tau^2_{\ell,k}(\boldsymbol{\zeta})] = \mathbb{E}[\tau^2_{\ell}(\boldsymbol{\zeta})] + \mathbb{E}[\tau^2_{k}(\boldsymbol{\zeta})] + \frac{\pi \int\limits_{-\infty}^{\infty} \omega^2 \mathcal{S}_{n_\ell n_\ell}(\omega)\mathcal{S}_{n_k n_k}(\omega)\mathrm{d}\omega}{T\alpha^2_\ell \alpha^2_k \big(\int\limits_{-\infty}^{\infty} \omega^2 \mathcal{S}_{ss}(\omega)\mathrm{d}\omega\big)^2}.$$

7. The equation for γ given in (5.2) was developed under the assumption that ToAs were estimated and the TDoAs were calculated from them as follows

$$\widehat{\mathsf{TDoA}}_{\ell,k}(\zeta) = \widehat{\mathsf{ToA}}_{\ell}(\zeta) - \widehat{\mathsf{ToA}}_{k}(\zeta).$$

That being the case, since $\widehat{\mathsf{ToA}}_{\ell}(\boldsymbol{\zeta})$ and $\widehat{\mathsf{ToA}}_{k}(\boldsymbol{\zeta})$ are independent, the variance for the TDoA estimator is

$$\mathbb{E}[\tau^2_{\ell,k}(\boldsymbol{\zeta})] = \mathbb{E}[\tau^2_{\ell}(\boldsymbol{\zeta})] + \mathbb{E}[\tau^2_{k}(\boldsymbol{\zeta})].$$

8. The variance of error on the TDoA estimates obtained directly from cross-correlation is greater than that from estimating the TDoAs as the difference between ToA estimates by the amount

$$\frac{\pi \int\limits_{-\infty}^{\infty} \omega^2 \mathcal{S}_{n_\ell n_\ell}(\omega)\mathcal{S}_{n_k n_k}(\omega)\mathrm{d}\omega}{T\alpha^2_\ell \alpha^2_k \big(\int\limits_{-\infty}^{\infty} \omega^2 \mathcal{S}_{ss}(\omega)\mathrm{d}\omega\big)^2}.$$

9. Parameters in (3.33) and (4.5) include the variances for the error on the times of arrival, but do not include the variances of the errors in the TDoAs. This means the equations do not take into account the extra noise term given in assertion 8.

10. Equations (3.33) and (4.5) with γ calculated using TDoA estimates in (5.2) will produce a good approximation for the *a posteriori* pdf if the noise term in assertion 8 is much smaller than every one of the L variances $\mathbb{E}[\tau^2_\ell(\boldsymbol{\zeta})]$.

11. From assertion 8 it is clear if only $(L-1)$ TDoA equations are going to be used to generate an *a posteriori* pdf, or even in a point estimator for that quality of the

approximate *a posteriori* pdf depends on the receiver used as the reference. The best reference would be the one that collectively minimized the $(L-1)$ extra variance terms.

In applications where the power spectral densities of the noise at all the receivers has the same shape, then choosing the receiver with the highest signal to noise ratio as the reference would produce the best estimate for the *a posteriori* pdf.

In conclusion the *a posteriori* pdf can be generated using (3.33) and (4.5) with γ computed by using TDoA measurements. However, the resulting pdf will only be the *a posteriori* if the TDoAs are obtained from the estimates of the ToAs. Using TDoAs estimated from cross-correlation will produce a pdf of some sort, it just won't be the *a posteriori*. It will be a good approximation to the *a posteriori* pdf if the signal to noise ratio of the (signal + noise) used a reference is significantly higher than the others.

A couple of observations, neither of which are related to the objective of this thesis, but worth pointing out nonetheless, are: (i) if the signal is known, the best results are obtained by using cross correlation to estimate the L ToAs rather than the set of $L(L-1)/2$ TDoAs, and (ii) no one set $(L-1)$ TDoA equations will contain all the statistical information available.

6. Summary and Suggestions for Further Studies

6.1 Summary

The contributions of this Ph.D. research are summarized as follows.

The first contribution is a novel localization method based on the theory of region estimation and Bayesian *a posteriori* pdf. Although this theory has been extensively explored, but has not been applied to the localization problem, at least not in the open literature. In particular, expressions for both 2D and 3D joint *a posteriori* pdfs of a target's coordinates were developed for time-of-arrivals of the source's signal in a line-of-sight environment, provided that locations of the gateways are measured. These expressions present approximations for the *a posteriori* pdfs that can also cope with the errors in the measurements of the gateways' locations. The approximations along with their bounds were developed based on the linearization of the relationship between the gateway's location error and the range to the target.

The obtained approximate *a posteriori* pdfs are functions of the measured times that a message arrives at spatially separated gateways, the distribution of the error corrupting the measurements of the ToAs and the gateways' locations, and the *a priori* pdf of the target. While the proposed method is applicable to any localization task that is based on time measurements, it was specifically presented in the context of geo-location in LoRaWAN. In particular the proposed method works with existing LoRaWAN infrastructure.

The proposed *a posteriori* pdfs of the target's coordinates were clearly positioned with respect to point estimators in literature. It was shown that the *a posteriori* pdfs of the target's coordinates can provide much more information than any point estimates. Particularly,

the *a posteriori* pdf is used to generate many types of point estimates as well as probabilities associated with any credible regions that are useful in practice.

In the second contribution, corroboration and utility of the *a posteriori* pdf were thoroughly developed for a wide range of practical scenarios.

Firstly, it was corroborated when errors in the measurements of gateways' locations are either present or not. In the absence of errors in the measurements of gateways' locations, the proposed *a posteriori* pdf becomes the exact pdf for the coordinates of a target and was corroborated by comparing with the approximate Gaussian *a posteriori* pdf obtained using Foy's linearized model. The comparison results revealed that the *a posteriori* pdf is not jointly Gaussian, especially when the target is near a gateway or when the *a priori* pdf is not Gaussian. This means that the Gaussian approximation to the *a posteriori* pdf cannot well reflect the effect of the gateway topology to the localization estimates. Therefore, whether 2D or 3D estimation is required, the *a posteriori* pdf should be preferred to the Gaussian approximation when the gateways are placed in the vicinity of each other.

The effects of errors of the measurements of the gateways' locations to the *a posteriori* pdf of the coordinates of a target were then explored. It was observed that the variance of such error contributes to the spread and the amount of this error causes a positional shift as well as contributes to the spread.

Secondly, the *a posteriori* pdf, which was initially developed for a single-transmission, was adapted to the scenario the target transmits multiple times from a fixed location to improve the quality of localization estimation. The *a posteriori* pdf for multiple transmissions was developed when the gateways' coordinates are measured either in every transmission or only once overtime.

In the first case, the *a posterior* pdf can be updated iteratively after each transmission since the measurements made for different transmission are independent. It is done by applying (3.33) iteratively with the *a priori* pdf being the *a posteriori* pdf calculated in the previous iteration.

In the second case, Equation (3.33) cannot be applied iteratively since the measured gateway coordinates are not independent from one transmission to the next. Instead, the averages of the ToA measurements in multiple transmissions are used as single measurements to calculate the *a posteriori* pdf using (3.33). Obviously, while the resulting *a posteriori* pdf is not the true one since it does not use all information in the measured ToAs from the multiple transmissions, it was shown to be an improvement over the *a posteriori* pdf obtained with a single transmission.

In the third contribution, the equation and approximation for the *a posteriori* pdf have been converted to accept TDoAs instead of ToAs, but the variances of the ToAs are still required. The equation is only valid if the TDoAs are obtained by taking the difference of ToAs since no information is lost in transforming L ToAs to $(L-1)$ TDoAs by subtracting one ToA from the other $(L-1)$ TDoAs. Furthermore, it does not matter which ToA is used as the reference.

If the TDoAs are measured directly through cross correlation, the variance of the error on the measurement error is greater than the sum of the variances of error on the times-of-arrival measured by correlation with a known signal.

If the TDoAs are measured directly through cross-correlation all $\frac{L(L-1)}{2}$ TDoA estimates carry some independent statistical information. It has been shown in [59] that linearly combining these TDoAs into a set of $(L-1)$ equations contains all the statistical information relevant to the ordinary least squares estimation, but the rendered $(L-1)$ TDoA estimates may not contain all of the statistical information and may not produce the *a posteriori* pdf.

6.2 Suggestions for Further Studies

Given that the methods and results presented in this thesis are for the LoS environment, it is of interest to consider the non-line-of-sight (NLOS) environment in future research. Also it would be interesting and useful to develop probabilistic localization that can makes use of multiple types of measurements (information), such as RSS and AoA, in addition to TDoA.

Appendix A

Consider the right side of (3.31) on page 46, which is $-\frac{1}{2}\sum_{\ell=1}^{L}\frac{(t-\mu_\ell)^2}{\sigma^2_{\mathsf{sum}_\ell}}$. Let a quadratic in t be given by

$$\begin{aligned} g(t) &= -\frac{1}{2}\sum_{\ell=1}^{L}\frac{(t-\mu_\ell)^2}{\sigma^2_{\mathsf{sum}_\ell}} \\ &= -\frac{1}{2}\sum_{\ell=1}^{L}\left(\frac{t^2}{\sigma^2_{\mathsf{sum}_\ell}} - \frac{2\mu_\ell t}{\sigma^2_{\mathsf{sum}_\ell}} + \frac{\mu_\ell^2}{\sigma^2_{\mathsf{sum}_\ell}}\right) \\ &= -\frac{1}{2}\left[\sum_{\ell=1}^{L}\frac{t^2}{\sigma^2_{\mathsf{sum}_\ell}} - 2\sum_{\ell=1}^{L}\frac{\mu_\ell t}{\sigma^2_{\mathsf{sum}_\ell}} + \sum_{\ell=1}^{L}\frac{\mu_\ell^2}{\sigma^2_{\mathsf{sum}_\ell}}\right]. \end{aligned}$$

Now let $\frac{1}{\sigma^2} = \sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}$. Then factor out $\frac{1}{\sigma^2}$ from the quadratic. This has

$$g(t) = -\frac{1}{2\sigma^2}\left[t^2 - 2\sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell t}{\sigma^2_{\mathsf{sum}_\ell}} + \sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell^2}{\sigma^2_{\mathsf{sum}_\ell}}\right].$$

Add $\left(\left(\sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2 - \left(\sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2\right)$, which is zero, into the square brackets. This allows $g(t)$ to be expressed as

$$\begin{aligned} g(t) &= -\frac{1}{2\sigma^2}\left[\left(t - \sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2 - \left(\sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2 + \sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell^2}{\sigma^2_{\mathsf{sum}_\ell}}\right] \\ &= -\frac{1}{2\sigma^2}\left[\left(t - \sigma^2\sum_{\ell=1}^{L}\frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2\right] - \frac{1}{2}\gamma, \end{aligned}$$

where $\gamma = \sum_{\ell=1}^{L}\frac{\mu_\ell^2}{\sigma^2_{\mathsf{sum}_\ell}} - \sigma^2\left(\sum_{\ell=1}^{L}\frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2$.

For the exponential of (3.31) on page 46

$$\mu = \sigma^2 \sum_{\ell=1}^{L} \frac{\mu_\ell}{\sigma^2_{\mathsf{sum}_\ell}},$$
$$\frac{1}{\sigma^2} = \sum_{\ell=1}^{L} \frac{1}{\sigma^2_{\mathsf{sum}_\ell}} = \sum_{\ell=1}^{L} \frac{1}{\sigma^2_{t_\ell} + \sigma^2_\ell / c^2},$$
$$\mu_\ell = \tau_\ell - \frac{\sqrt{(\widehat{x}_\ell - x)^2 + (\widehat{y}_\ell - y)^2}}{c}.$$

Appendix B

The likelihood function for the case that $t(\zeta)$ is known is obtained by the density function for $t(\boldsymbol{\zeta})$ being Gaussian with a mean of μ equal to the value of known t and a variance that tends to zero. After integration with respect to t, the likelihood function is the numerator in the right side of (3.30). It is an exponential with the exponent given by (3.31) on page 46 with the variable t now being its known value. The first derivative of $-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\left(t-\mu_\ell\right)^2}{\sigma^2_{\mathsf{sum}_\ell}}$ given on page 46 is

$$\begin{aligned}
\frac{\partial\left(-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\left(t-\mu_\ell\right)^2}{\sigma^2_{\mathsf{sum}_\ell}}\right)}{\partial x} &= -\sum_{\ell=1}^{L}\frac{\left(t-\tau_\ell+\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c\right)(x-\widehat{x}_\ell)}{\sigma^2_{\mathsf{sum}_\ell}c\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}},\\
\frac{\partial\left(-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\left(t-\mu_\ell\right)^2}{\sigma^2_{\mathsf{sum}_\ell}}\right)}{\partial y} &= -\sum_{\ell=1}^{L}\frac{\left(t-\tau_\ell+\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c\right)(y-\widehat{y}_\ell)}{\sigma^2_{\mathsf{sum}_\ell}c\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}}.
\end{aligned} \tag{B.1}$$

The first derivative of γ given on page 47 is

$$\begin{aligned}
\frac{\partial\gamma}{\partial x} =& -\sum_{\ell=1}^{L}\frac{2\big(\tau_\ell-\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c\big)(x-\widehat{x}_\ell)}{\sigma^2_{\mathsf{sum}_\ell}\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}}\\
&+2\left(\frac{1}{\sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}}\right)\left(\sum_{\ell=1}^{L}\frac{\tau_\ell-\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c}{\sigma^2_{\mathsf{sum}_\ell}}\right)\\
&\left(\sum_{\ell=1}^{L}\frac{x-\widehat{x}_\ell}{\sigma^2_{\mathsf{sum}_\ell}\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}}\right),\\
\frac{\partial\gamma}{\partial y} =& -\sum_{\ell=1}^{L}\frac{2\big(\tau_\ell-\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c\big)(y-\widehat{y}_\ell)}{\sigma^2_{\mathsf{sum}_\ell}\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}}\\
&+2\left(\frac{1}{\sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}}\right)\left(\sum_{\ell=1}^{L}\frac{\tau_\ell-\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c}{\sigma^2_{\mathsf{sum}_\ell}}\right)\\
&\left(\sum_{\ell=1}^{L}\frac{y-\widehat{y}_\ell}{\sigma^2_{\mathsf{sum}_\ell}\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}}\right).
\end{aligned} \tag{B.2}$$

Appendix C

Without loss of generality, let τ_1 be the reference to calculate the time-difference-of-arrivals. The left side of (3.31) on page 46 is rewritten as

$$-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\big((t+\tau_1)-(\tau_\ell-\tau_1-\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c)\big)^2}{\sigma^2_{\mathsf{sum}_\ell}}.$$

Let the quadratic in $(t+\tau_1)$ be written as

$$g(t+\tau_1)=-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\big((t+\tau_1)-(\tau_\ell-\tau_1-\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c)\big)^2}{\sigma^2_{\mathsf{sum}_\ell}}. \tag{C.1}$$

Replacing $T=t+\tau_1$, $\tau_{\ell,1}=\tau_\ell-\tau_1$ and $\widehat{r}_\ell=\sqrt{(\widehat{x}_\ell-x)^2+(\widehat{y}_\ell-y)^2}/c$ allows to simplify (C.1) to

$$\begin{aligned}
g(T)&=-\frac{1}{2}\sum_{\ell=1}^{L}\frac{\big(T-(\tau_{\ell,1}-\widehat{r}_\ell)\big)^2}{\sigma^2_{\mathsf{sum}_\ell}}\\
&=-\frac{1}{2}\sum_{\ell=1}^{L}\left(\frac{T^2}{\sigma^2_{\mathsf{sum}_\ell}}-\frac{2\big(\tau_{\ell,1}-\widehat{r}_\ell\big)T}{\sigma^2_{\mathsf{sum}_\ell}}+\frac{\big(\tau_{\ell,1}-\widehat{r}_\ell\big)^2}{\sigma^2_{\mathsf{sum}_\ell}}\right)\\
&=-\frac{1}{2}\left[\sum_{\ell=1}^{L}\frac{T^2}{\sigma^2_{\mathsf{sum}_\ell}}-2\sum_{\ell=1}^{L}\frac{\big(\tau_{\ell,1}-\widehat{r}_\ell\big)T}{\sigma^2_{\mathsf{sum}_\ell}}+\sum_{\ell=1}^{L}\frac{\big(\tau_{\ell,1}-\widehat{r}_\ell\big)^2}{\sigma^2_{\mathsf{sum}_\ell}}\right].
\end{aligned}$$

Now let $\frac{1}{\sigma^2}=\sum_{\ell=1}^{L}\frac{1}{\sigma^2_{\mathsf{sum}_\ell}}$. Then factoring out $\frac{1}{\sigma^2}$ from the quadratic has

$$g(T)=-\frac{1}{2\sigma^2}\left[T^2-2\sigma^2\sum_{\ell=1}^{L}\frac{\big(\tau_{\ell,1}-\widehat{r}_\ell\big)T}{\sigma^2_{\mathsf{sum}_\ell}}+\sigma^2\sum_{\ell=1}^{L}\frac{\big(\tau_{\ell,1}-\widehat{r}_\ell\big)^2}{\sigma^2_{\mathsf{sum}_\ell}}\right].$$

Add $\left(\left(\sigma^2\sum_{\ell=1}^{L}\frac{(\tau_{\ell,1}-\widehat{r}_\ell)}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2-\left(\sigma^2\sum_{\ell=1}^{L}\frac{(\tau_{\ell,1}-\widehat{r}_\ell)}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2\right)$, which is zero, into the square brackets. This allows $g(T)$ to be expressed as

$$
\begin{aligned}
g(T) &= -\frac{1}{2\sigma^2}\left[\left(T-\sigma^2\sum_{\ell=1}^{L}\frac{\left(\tau_{\ell,1}-\widehat{r}_\ell\right)}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2-\left(\sigma^2\sum_{\ell=1}^{L}\frac{\left(\tau_{\ell,1}-\widehat{r}_\ell\right)}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2+\sigma^2\sum_{\ell=1}^{L}\frac{\left(\tau_{\ell,1}-\widehat{r}_\ell\right)^2}{\sigma^2_{\mathsf{sum}_\ell}}\right] \\
&= -\frac{1}{2\sigma^2}\left[\left(T-\sigma^2\sum_{\ell=1}^{L}\frac{\left(\tau_{\ell,1}-\widehat{r}_\ell\right)}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2\right]-\frac{1}{2}\gamma,
\end{aligned}
$$

where

$$
\gamma=\sum_{\ell=1}^{L}\frac{\left(\tau_{\ell,1}-\widehat{r}_\ell\right)^2}{\sigma^2_{\mathsf{sum}_\ell}}-\sigma^2\left(\sum_{\ell=1}^{L}\frac{\left(\tau_{\ell,1}-\widehat{r}_\ell\right)}{\sigma^2_{\mathsf{sum}_\ell}}\right)^2. \tag{C.2}
$$

Equation (C.2) shows that the *a posteriori* pdf given by Equations (3.33) and (4.5) can be generated using TDoA values that are obtained from the differences of the ToA measurements.

Appendix D

This appendix finds a bound on the error after linearizing the range with respect to the measurement error on the coordinates of the gateways. The actual range from gateway ℓ is

$$r_\ell(\boldsymbol{\zeta}) = \sqrt{(x(\boldsymbol{\zeta}) - x_\ell)^2 + (y(\boldsymbol{\zeta}) - y_\ell)^2}.$$

Substituting $\widehat{x}_\ell(\boldsymbol{\zeta}) - \Delta x_\ell(\boldsymbol{\zeta})$ for x_ℓ and $\widehat{y}_\ell(\boldsymbol{\zeta}) - \Delta y_\ell(\boldsymbol{\zeta})$ for y_ℓ has

$$r_\ell(\boldsymbol{\zeta}) = \sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + \Delta x_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}) + \Delta y_\ell(\boldsymbol{\zeta})\big)^2}.$$

The equation is linearized in $\Delta x_\ell(\boldsymbol{\zeta})$ and $\Delta y_\ell(\boldsymbol{\zeta})$ with a first order Maclaurin polynomial to get the approximation

$$\begin{aligned} r_\ell(\boldsymbol{\zeta}) \simeq \sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\big)^2} + \left(\frac{\partial r_\ell(\boldsymbol{\zeta})}{\partial \Delta x_\ell(\boldsymbol{\zeta})}\bigg|_{\Delta x_\ell(\boldsymbol{\zeta})=0, \Delta y_\ell(\boldsymbol{\zeta})=0} \right) \Delta x_\ell(\boldsymbol{\zeta}) \\ + \left(\frac{\partial r_\ell(\boldsymbol{\zeta})}{\partial \Delta y_\ell(\boldsymbol{\zeta})}\bigg|_{\Delta x_\ell(\boldsymbol{\zeta})=0, \Delta y_\ell(\boldsymbol{\zeta})=0} \right) \Delta y_\ell(\boldsymbol{\zeta}), \end{aligned} \tag{D.1}$$

where

$$\frac{\partial r_\ell(\boldsymbol{\zeta})}{\partial \Delta x_\ell(\boldsymbol{\zeta})} = \frac{x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + \Delta x_\ell(\boldsymbol{\zeta})}{\sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + \Delta x_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}) + \Delta y_\ell(\boldsymbol{\zeta})\big)^2}}.$$

Evaluating the above at $\Delta x_\ell(\boldsymbol{\zeta}) = 0$ and $\Delta y_\ell(\boldsymbol{\zeta}) = 0$ and substituting the result into (D.1) yields the Maclaurin polynomial approximation

$$\begin{aligned} r_\ell(\boldsymbol{\zeta}) \approx & \sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\big)^2} + \frac{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))\Delta x_\ell(\boldsymbol{\zeta})}{\sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\big)^2}} \\ & + \frac{(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))\Delta y_\ell(\boldsymbol{\zeta})}{\sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\big)^2}}. \end{aligned} \tag{D.2}$$

The error, i.e., the remainder, expressed in Lagrange form is given by

$$\mathsf{P}_{r_\ell}(z_1, z_2) = \left(\frac{\partial^2 r_\ell(\boldsymbol{\zeta})}{\partial \Delta x_\ell^2(\boldsymbol{\zeta})}\bigg|_{\Delta x_\ell(\boldsymbol{\zeta})=z_1} \right) \frac{\Delta x_\ell^2(\boldsymbol{\zeta})}{2} + \left(\frac{\partial^2 r_\ell(\boldsymbol{\zeta})}{\partial \Delta y_\ell^2(\boldsymbol{\zeta})}\bigg|_{\Delta y_\ell(\boldsymbol{\zeta})=z_2} \right) \frac{\Delta y_\ell^2(\zeta)}{2}$$

for some value of z_1 between 0 and $\Delta x_\ell(\boldsymbol{\zeta})$ and some value of z_2 between 0 and $\Delta y_\ell(\boldsymbol{\zeta})$, where

$$\begin{aligned} \frac{\partial^2 r_\ell(\boldsymbol{\zeta})}{\partial \Delta x_\ell^2(\boldsymbol{\zeta})}\bigg|_{\Delta x_\ell(\boldsymbol{\zeta})=z_1, \Delta y_\ell(\boldsymbol{\zeta})=0} &= -\frac{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))^2}{\left(\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + z_1)^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}\right)^3} \\ &\quad + \frac{1}{\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + z_1)^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}} \\ &= \frac{1 - \frac{\left(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\right)^2}{\left(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + z_1\right)^2 + \left(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\right)^2}}{\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}) + z_1)^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}}. \end{aligned} \tag{D.3}$$

Similarly,

$$\frac{\partial^2 r_\ell(\boldsymbol{\zeta})}{\partial \Delta y_\ell^2(\boldsymbol{\zeta})}\bigg|_{\Delta x_\ell(\boldsymbol{\zeta})=0,\, \Delta y_\ell(\boldsymbol{\zeta})=z_2} = \frac{1 - \frac{\left(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\right)^2}{\left(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\right)^2 + \left(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}) + z_2\right)^2}}{\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}) + z_2)^2}}. \tag{D.4}$$

The error in the approximation given by (D.2) is certain to be less than the maximum of $\left|\mathsf{P}_{r_\ell}(z_1, z_2)\right|$ over the domain $\left|z_1\right| \le \left|\Delta x_\ell(\boldsymbol{\zeta})\right|,\ \left|z_2\right| \le \left|\Delta y_\ell(\boldsymbol{\zeta})\right|$. Unfortunately, the maximum depends on the values of $(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))$ and $(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))$ and could be ∞, which would render it useless as an upper bound. However, under the assumption the target is spatially separated from each gateway ℓ such that $|x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})| > (2+\sqrt{2})|\Delta x_\ell(\boldsymbol{\zeta})|$ and $|y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})| > (2+\sqrt{2})|\Delta y_\ell(\boldsymbol{\zeta})|$, the error is reasonably bounded. The numerators on the right sides of (D.3) and (D.4) are bounded below by 0 and above by 1 for z_1 between $\pm\Delta x_\ell(\boldsymbol{\zeta})$ and z_2 between $\pm\Delta y_\ell(\boldsymbol{\zeta})$.

Replacing the numerators on the right sides of (D.3) and (D.4) with 1 and recognizing

the resulting expressions is maximized when $z_1 = 0$ and $z_2 = 0$ produces the inequalities

$$\max_{z_1,\, z_2} \left| \frac{\partial^2 r_\ell(\boldsymbol{\zeta})}{\partial \Delta x_\ell^2(\boldsymbol{\zeta})} \right|_{\substack{\Delta x_\ell(\boldsymbol{\zeta})=z_1,\\ \Delta y_\ell(\boldsymbol{\zeta})=0}} \leq \frac{1}{\sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\big)^2}},$$

and

$$\max_{z_1,\, z_2} \left| \frac{\partial^2 r_\ell(\boldsymbol{\zeta})}{\partial \Delta y_\ell^2(\boldsymbol{\zeta})} \right|_{\substack{\Delta x_\ell(\boldsymbol{\zeta})=0,\\ \Delta y_\ell(\boldsymbol{\zeta})=z_2}} \leq \frac{1}{\sqrt{\big(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})\big)^2 + \big(y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})\big)^2}}$$

which are certain to be true if the separation between the target and the measured x and y coordinates of gateway ℓ are such that $|x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})| > (2+\sqrt{2})|\Delta x_\ell(\boldsymbol{\zeta})|$ and $|y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})| > (2+\sqrt{2})|\Delta y_\ell(\boldsymbol{\zeta})|$.

Therefore, the errors in the second and third terms of (D.2), which are due to measurement errors $\Delta x_\ell(\boldsymbol{\zeta})$ and $\Delta y_\ell(\boldsymbol{\zeta})$, respectively, are bounded by

$$\big|e_{\Delta x_\ell}\big| \leq \frac{\Delta x_\ell(\boldsymbol{\zeta})^2}{2\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}},$$

and

$$\big|e_{\Delta y_\ell}\big| \leq \frac{\Delta y_\ell(\boldsymbol{\zeta})^2}{2\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}}.$$

The absolute error is not as important as the relative error in the second and third terms on the right side of (D.2). The ratios of $\big|e_{\Delta x_\ell}\big|$ and $\big|e_{\Delta x_\ell}\big|$ to the second and third terms on the right side of (D.2), respectively, are

$$\begin{aligned} \frac{\big|e_{\Delta x_\ell}\big|}{\text{Second Term}} &\leq \frac{\frac{\Delta x_\ell(\boldsymbol{\zeta})^2}{2\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}}}{\frac{|(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))\Delta x_\ell(\boldsymbol{\zeta})|}{\sqrt{(x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta}))^2 + (y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta}))^2}}} \\ &\leq \frac{|\Delta x_\ell(\boldsymbol{\zeta})|}{2|x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})|}, \end{aligned}$$

and

$$\frac{\left|e_{\Delta y_\ell}\right|}{\text{Third Term}} \leq \frac{|\Delta y_\ell(\boldsymbol{\zeta})|}{2|y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})|}$$

providing the target is spaced from all L the gateways such that $|x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})| > (2 + \sqrt{2})|\Delta x_\ell(\boldsymbol{\zeta})|$ and $|x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})| > (2 + \sqrt{2})|\Delta x_\ell(\boldsymbol{\zeta})|$.

This means, for example, if $|x(\boldsymbol{\zeta}) - \widehat{x}_\ell(\boldsymbol{\zeta})| > 5|\Delta x_\ell(\boldsymbol{\zeta})|$ and $|y(\boldsymbol{\zeta}) - \widehat{y}_\ell(\boldsymbol{\zeta})| > 5|\Delta y_\ell(\boldsymbol{\zeta})|$, then the components of (D.2) due to Δx_ℓ and Δy_ℓ will be in error by at most 10%.

Appendix E

The Gaussian Approximation to the *a Posteriori* Density

An approximation to the *a posteriori* density is obtained by linearizing the set of non-linear equations about their least squared error solution. The linearized model allows the covariances of the co-ordinates of the target's location to be computed from the covariances of the times-of-arrival in units of m^2. The logical progression from the set of non-linear equations to the covariance matrix for the co-ordinate estimators is given below. The set of non-linear range equations for outcome ζ is

$$r_\ell(\zeta) = c(\widehat{\mathsf{ToA}}_\ell(\zeta) - t(\zeta)), \quad \ell = 1, 2, \ldots, L,$$

where c is the speed of light, $r_\ell(\zeta)$ is the range from the target to gateway ℓ and is given by

$$r_\ell(\zeta) = \sqrt{(x(\zeta) - x_\ell)^2 + (y(\zeta) - y_\ell)^2 + (z(\zeta) - z_\ell)^2},$$

$\widehat{\mathsf{ToA}}_\ell(\zeta)$ is the measured time that the message arrives at gateway ℓ and $t(\zeta)$ is the time the message is transmitted.

The set of non-linear equations are re-organized to put the quantities that are measured on the right side and the quantities being estimated on the left side with the result, again for outcome ζ, being

$$r_\ell(\zeta) + ct(\boldsymbol{\zeta}) = c\widehat{\mathsf{ToA}}_\ell(\zeta), \quad \ell = 1, 2, \ldots, L.$$

Then this set of equations is re-arranged into another set, which is the generalized least squared form, by weighting and combining the equations with weighting factors $w_{i,k}$. The

new set of L generalized non-linear equations are expressed in the matrix form by

$$W \begin{bmatrix} r_1(\zeta) + ct(\zeta) \\ r_2(\zeta) + ct(\zeta) \\ \vdots \\ r_L(\zeta) + ct(\zeta) \end{bmatrix} = W \begin{bmatrix} c\widehat{\mathsf{ToA}}_1(\zeta) \\ c\widehat{\mathsf{ToA}}_2(\zeta) \\ \vdots \\ c\widehat{\mathsf{ToA}}_L(\zeta) \end{bmatrix}, \tag{E.1}$$

where W is the $L \times L$ matrix with elements $w_{i,k}$.

It is well known the squared error is minimized when $W = \sqrt{C^{-1}}$ (see A. C. Aitken, 1934), where C is the positive definite covariance matrix for time-of-arrival vector

$$\left[c\widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}), c\widehat{\mathsf{ToA}}_2(\boldsymbol{\zeta}), \ldots, c\widehat{\mathsf{ToA}}_L(\boldsymbol{\zeta})\right].$$

The elements of C are co-variances in unit of m^2.

Suppose the non-linear least squares solutions to the generalized non-linear equations are $\widehat{x}$, $\widehat{y}$, $\widehat{z}$ and $\widehat{t}$. Then the approximation to the *a posteriori* density is made under the following assumptions:

- The target is located at $\widehat{x}$, $\widehat{y}$, $\widehat{z}$ and the time the message was transmitted is $\widehat{t}$. Of course this is unlikely to be true, but that is the assumption. However, the target should be in the vicinity.
- The errors in the time-of-arrival measurements are so small that the measurement errors in the ToAs and the estimation errors in the co-ordinates of the target are accurately related by the first order Taylor series expansion of the non-linear equations about the point $(x, y, z) = (\widehat{x}, \widehat{y}, \widehat{z})$ and $t = \widehat{t}$.

Under these two assumptions, the errors in the time-of-arrival measurements, which are random variables that will be denoted $\Delta\widehat{\mathsf{ToA}}_\ell(\boldsymbol{\zeta})$ for gateway ℓ, and the errors in the estimators that will be denoted $\Delta x(\boldsymbol{\zeta})$, $\Delta y(\boldsymbol{\zeta})$, $\Delta z(\boldsymbol{\zeta})$ and $\Delta t(\boldsymbol{\zeta})$, are related by the set of

linear equations.

$$WA \begin{bmatrix} \Delta x(\boldsymbol{\zeta}) \\ \Delta y(\boldsymbol{\zeta}) \\ \Delta z(\boldsymbol{\zeta}) \\ c\Delta t(\boldsymbol{\zeta}) \end{bmatrix} = W \begin{bmatrix} c\Delta\widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}) \\ c\Delta\widehat{\mathsf{ToA}}_2(\boldsymbol{\zeta}) \\ \vdots \\ c\Delta\widehat{\mathsf{ToA}}_L(\boldsymbol{\zeta}) \end{bmatrix}, \tag{E.2}$$

where A is an $L \times 4$ matrix of the coefficients for the first order Taylor series expansion about the points $(x(\zeta), y(\zeta), x(\zeta)) = (\widehat{x},\, \widehat{y},\, \widehat{z})$ and $t(\zeta) = \widehat{t}$. The matrix A is given by

$$\begin{aligned} A &= \begin{bmatrix} \frac{\partial r_1(\zeta)}{\partial x} & \frac{\partial r_1(\zeta)}{\partial y} & \frac{\partial r_1(\zeta)}{\partial z} & \frac{\partial ct(\zeta)}{c\partial t} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial r_L(\zeta)}{\partial x} & \frac{\partial r_L(\zeta)}{\partial y} & \frac{\partial r_L(\zeta)}{\partial z} & \frac{\partial ct(\zeta)}{c\partial t} \end{bmatrix}_{(x(\zeta),y(\zeta),x(\zeta))=(\widehat{x},\widehat{y},\widehat{z}),\, t(\zeta)=\widehat{t}} \\ &= \begin{bmatrix} \frac{\widehat{x}-x_1}{\widehat{r}_1} & \frac{\widehat{y}-y_1}{\widehat{r}_1} & \frac{\widehat{z}-z_1}{\widehat{r}_1} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\widehat{x}-x_L}{\widehat{r}_L} & \frac{\widehat{y}-y_L}{\widehat{r}_L} & \frac{\widehat{z}-z_L}{\widehat{r}_L} & 1 \end{bmatrix}, \end{aligned} \tag{E.3}$$

where $\widehat{r}_\ell = \sqrt{(\widehat{x} - x_\ell)^2 + (\widehat{y} - y_\ell)^2 + (\widehat{z} - z_\ell)^2}$. The generalized least squared error solution yields estimator errors $\Delta x(\boldsymbol{\zeta})$, $\Delta y(\boldsymbol{\zeta})$, $\Delta z(\boldsymbol{\zeta})$ and $c\Delta t(\boldsymbol{\zeta})$ given by

$$\begin{bmatrix} \Delta x(\boldsymbol{\zeta}) \\ \Delta y(\boldsymbol{\zeta}) \\ \Delta z(\boldsymbol{\zeta}) \\ c\Delta t(\boldsymbol{\zeta}) \end{bmatrix} = (A^\top W^\top W A)^{-1} A^\top W^\top W \begin{bmatrix} c\Delta\widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}) \\ c\Delta\widehat{\mathsf{ToA}}_2(\boldsymbol{\zeta}) \\ \vdots \\ c\Delta\widehat{\mathsf{ToA}}_L(\boldsymbol{\zeta}) \end{bmatrix}. \tag{E.4}$$

The least squared error solution with the optimum weighting matrix $W = \sqrt{C^{-1}}$ reduces

to

$$\begin{bmatrix} \Delta x(\boldsymbol{\zeta}) \\ \Delta y(\boldsymbol{\zeta}) \\ \Delta z(\boldsymbol{\zeta}) \\ c\Delta t(\boldsymbol{\zeta}) \end{bmatrix} = (A^\top C^{-1} A)^{-1} A^\top C^{-1} \begin{bmatrix} c\Delta\widehat{\mathsf{ToA}}_1(\boldsymbol{\zeta}) \\ c\Delta\widehat{\mathsf{ToA}}_2(\boldsymbol{\zeta}) \\ \vdots \\ c\Delta\widehat{\mathsf{ToA}}_L(\boldsymbol{\zeta}) \end{bmatrix}, \tag{E.5}$$

Since C is positive definite, so is C^{-1}. Therefore, C^{-1} can be factored into the product of two positive definite matrices $C^{-1} = \sqrt{C^{-1}}\sqrt{C^{-1}}$. With $W = \sqrt{C^{-1}}$, the covariance matrix for the estimators is given by

$$\begin{aligned} &\mathbb{E}\left[[\Delta x(\boldsymbol{\zeta}), \Delta y(\boldsymbol{\zeta}), \Delta z(\boldsymbol{\zeta}), \Delta t(\boldsymbol{\zeta})]^\top [\Delta x(\boldsymbol{\zeta}), \Delta y(\boldsymbol{\zeta}), \Delta z(\boldsymbol{\zeta}), \Delta t(\boldsymbol{\zeta})]\right] \\ &\qquad = \big((A^\top C^{-1} A)^{-1} A^\top C^{-1}\big) C \big((A^\top C^{-1} A)^{-1} A^\top C^{-1}\big)^\top \\ &\qquad = (A^\top C^{-1} A)^{-1} \end{aligned} \tag{E.6}$$

The covariance matrix for the estimators is therefore given by

$$\begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix} = (A^\top C^{-1} A)^{-1}. \tag{E.7}$$

If the measurement errors, i.e., $c\Delta\mathsf{ToA}_\ell(\boldsymbol{\zeta}), \quad \ell = 1, 2, \ldots, L$ are jointly Gaussian, then the estimation errors will also be jointly Gaussian. The conditional pdf is therefore given by

$$\begin{aligned} &f_{\Delta x, \Delta y, \Delta z, c\Delta t \,|\, \widehat{\mathsf{ToA}}}(\Delta x, \Delta y, \Delta z, c\Delta t \,|\, \boldsymbol{\tau}) = \\ &\qquad \frac{\left|A^\top C^{-1} A\right| \exp(-\frac{1}{2}[\Delta x, \Delta y, \Delta z, c\Delta t] A^\top C^{-1} A [\Delta x, \Delta y, \Delta z, c\Delta t]^\top)}{4\pi^2}, \end{aligned} \tag{E.8}$$

This joint pdf is the approximate *a posteriori* pdf.

It must be emphasized that this approximate joint conditional density, which is an approx-

imation to the *a posteriori* density, was found without any numerical integration. Further more, the marginal densities are Gaussian with variances σ_x^2, σ_y^2 and σ_z^2.

Dilution of Precision

In the special case where $\Delta\mathsf{ToA}_\ell(\boldsymbol{\zeta}), \quad \ell = 1, 2, \ldots, L$ are independent with identical variance σ^2, then the covariance matrix is

$$C_{\mathsf{ToA}} = \mathbb{E}\left[c[\widehat{\mathsf{ToA}}_1, \widehat{\mathsf{ToA}}_2, \ldots, \widehat{\mathsf{ToA}}_L]^\top \times c[\widehat{\mathsf{ToA}}_1, \widehat{\mathsf{ToA}}_2, \ldots, \widehat{\mathsf{ToA}}_L]\right]$$

a diagonal matrix with all diagonal entries being $\sigma_m^2 = c^2\sigma^2$, where σ_m^2 has units m^2. In this special case the covariance of the estimation errors is given by

$$C_{\mathsf{EST}} \equiv \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} & \sigma_{xt} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} & \sigma_{yt} \\ \sigma_{zx} & \sigma_{zy} & \sigma_z^2 & \sigma_{zt} \\ \sigma_{tx} & \sigma_{ty} & \sigma_{tz} & \sigma_t^2 \end{bmatrix} = \sigma_m^2 (A^\top A)^{-1}, \tag{E.9}$$

where all covariances have units m^2. Dividing the equation by the scalar σ_m^2 has

$$(A^\top A)^{-1} = \begin{bmatrix} \frac{\sigma_x^2}{\sigma_m^2} & \frac{\sigma_{xy}}{\sigma_m^2} & \frac{\sigma_{xz}}{\sigma_m^2} & \frac{\sigma_{xt}}{\sigma_m^2} \\ \frac{\sigma_{yx}}{\sigma_m^2} & \frac{\sigma_y^2}{\sigma_m^2} & \frac{\sigma_{yz}}{\sigma_m^2} & \frac{\sigma_{yt}}{\sigma_m^2} \\ \frac{\sigma_{zx}}{\sigma_m^2} & \frac{\sigma_{zy}}{\sigma_m^2} & \frac{\sigma_z^2}{\sigma_m^2} & \frac{\sigma_{zt}}{\sigma_m^2} \\ \frac{\sigma_{tx}}{\sigma_m^2} & \frac{\sigma_{ty}}{\sigma_m^2} & \frac{\sigma_{tz}}{\sigma_m^2} & \frac{\sigma_t^2}{\sigma_m^2} \end{bmatrix}. \tag{E.10}$$

In this scenario the diagonal elements of $(A^\top A)^{-1}$ are the ratio of the variances of the estimates to the variances of the measurement error. When the covariances are represented as a ratio, the ratio becomes a "magnification factor" from the variance of the time-of-arrival (in m^2) to the variance in x, y or z (in m^2). The "magnification factor" is the A matrix, which in turn is entirely dependent on the spatial relationship among the gateways and target.

The square root of the trace of $(A^\top A)^{-1}$, i.e., the sum square root of the diagonal elements, is called the geometric dilution of precision (GDoP), the quantity $\sqrt{\frac{\sigma_x^2}{\sigma_m^2}+\frac{\sigma_y^2}{\sigma_m^2}+\frac{\sigma_z^2}{\sigma_m^2}}$ is called the position dilution of precision (PDoP), the quantity $\sqrt{\frac{\sigma_x^2}{\sigma_m^2}+\frac{\sigma_y^2}{\sigma_m^2}}$ is called the horizontal dilution of precision (HDoP), and $\sqrt{\frac{\sigma_z^2}{\sigma_m^2}}$ is called the vertical dilution of precision (VDoP).

To give context to the accuracy of the Gaussian approximation to the *a posteriori* joint density, the Gaussian approximation will be compared to the actual *a posteriori* marginal densities graphically.

However, before doing that some general comments on the accuracy are mentioned.

There are situations were the approximate 3 dimensional *a posteriori* density cannot be found. Even if there are 4 or more gateways in view of the target. This situation arises for the system of Figure 4.1 on page 57 when the target is located on either the x or y-axis and the four gateways are mounted at the same elevation.

There are also situations where the *a posteriori* pdf is not jointly Gaussian. Of course, the approximate *a posteriori* density, by the way it was derived, must be jointly Gaussian. Therefore, the approximate pdf will not be a good approximation in these situations.

In general, the 3D Gaussian approximation to the approximate *a posteriori* pdf is not a great fit to the 3D *a posteriori* pdf. However, under conditions where the z co-ordinate is known or assumed, the Gaussian approximation to the 2D *a posteriori*, which is found by including $(z - z_\ell)$ in the r_ℓ, but removing the third column in A, is a good fit pdf providing the 2D *a posteriori* pdf is Gaussian.

The situations pointed out above, except for the case where the *a priori* pdf is not uniform, are illustrated with examples. In all examples the *a posteriori* and the approximation are generated using the "very lucky" observation where there is no measurement error. The error on the time-of-arrival measurement is assumed to be Gaussian with a standard deviation of 25ns (7.5m). The errors were assumed to be i.i.d so the covariance matrix for

$c\widehat{\mathsf{ToA}}(\boldsymbol{\zeta})$ was

$$C_{\mathsf{ToA}} = \begin{bmatrix} \sigma_m^2 & 0 & 0 & 0 \\ 0 & \sigma_m^2 & 0 & 0 \\ 0 & 0 & \sigma_m^2 & 0 \\ 0 & 0 & 0 & \sigma_m^2 \end{bmatrix} \tag{E.11}$$

where $\sigma_m^2 = (7.5m)^2$. This is somewhat optimistic value for σ_m, which is realistically 25 to 30m (85 to 100ns). The smaller σ_m is favorable to the linearized model used in the approximation. Thus the Gaussian approximation in the examples will be better than they would in practice.

The examples are based on the scenarios of Figure 4.1 on page 57, with the gateways either all located at 25m or located at elevations $[z_1,\ z_2,\ z_3,\ z_4] = [25,\ 225,\ 425,\ 625]$m.

Incorporating *a Priori* Information in the Gaussian Approximation

A priori information can also be incorporated into the Gaussian approximation to the *a posteriori* probability density function. The Gaussian approximation was developed under the assumption that the *a priori* density was uniform (i.e., constant). However, the *a priori* information can be easily integrated into the Gaussian approximation, but, unless the *a priori* density is jointly Gaussian, the computation of the approximate *a posteriori* pdf will require a triple numerical integration for 3D estimation and double numerical integration for 2D estimation. Should numerical integration be required to transfer the information contained in the *a priori* pdf into the *a posteriori* pdf, then the primary advantage of the Gaussian approximation, which is its computational efficiency, is lost.

The reasoning and methodology for transferring *a priori* information into the Gaussian approximation of the *a posteriori* pdf is outlined below.

First, it is necessary to understand how the information in the *a priori* pdf is transferred

into the *a posteriori* pdf with Bayes equation. Bayes equation states:

$$f_{x,y,z,t\left|\widehat{\mathsf{ToA}}\right.}(x,y,z,t\left|\boldsymbol{\tau}\right.) = \frac{f_{\widehat{\mathsf{ToA}}\left|_{x,y,z,t}\right.}(\boldsymbol{\tau}\left|x,y,z,t\right.)f_{x,y,z,t}(x,y,z,t)}{\mathsf{volume}_1},$$

where

$$\mathsf{volume}_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathsf{f}_{\widehat{\mathsf{ToA}}\left|_{\mathsf{x,y,z,t}}\right.}(\boldsymbol{\tau}\left|\mathsf{x,y,z,t}\right.)\mathsf{f}_{\mathsf{x,y,z,t}}(\mathsf{x,y,z,t})\mathrm{dtdxdydz}.$$

The joint *a posteriori* pdf for just the co-ordinates is obtained by integrating w.r.t t as indicated below

$$\int_{-\infty}^{\infty} f_{x,y,z,t\left|\widehat{\mathsf{ToA}}\right.}(x,y,z,t\left|\boldsymbol{\tau}\right.)\mathrm{d}t = \frac{\int_{-\infty}^{\infty} f_{\widehat{\mathsf{ToA}}\left|_{x,y,z,t}\right.}(\boldsymbol{\tau}\left|x,y,z,t\right.)f_{x,y,z,t}(x,y,z,t)\mathrm{d}t}{\mathsf{volume}_1} \tag{E.12}$$

If $f_{t|x,y,z}(t|x,y,z)$ is widely uniform (i.e., constant) and does not depend on x, y or z, then

$$\begin{aligned} f_{x,y,z,t}(x,y,z,t) &= f_{t|x,y,z}(t|x,y,z)f_{x,y,z}(x,y,z) \\ &= \mathsf{constant} \times f_{x,y,z}(x,y,z). \end{aligned}$$

Making the substitution above into (E.12) shows the integral w.r.t t in both the numerator and denominator can be evaluated symbolically regardless of the shape of $f_{x,y,z}(x,y,z)$. After performing the integrals, (E.12) can be written as

$$f_{x,y,z\left|\widehat{\mathsf{ToA}}\right.}(x,y,z\left|\boldsymbol{\tau}\right.) = \frac{f_{\widehat{\mathsf{ToA}}\left|_{x,y,z}\right.}(\boldsymbol{\tau}\left|x,y,z\right.)f_{x,y,z}(x,y,z)}{\mathsf{volume}_1}, \tag{E.13}$$

where

$$\mathsf{volume}_1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathsf{f}_{\widehat{\mathsf{ToA}}\left|_{\mathsf{x,y,z}}\right.}(\boldsymbol{\tau}\left|\mathsf{x,y,z}\right.)\mathsf{f}_{\mathsf{x,y,z}}(\mathsf{x,y,z})\mathrm{dxdydz}.$$

Note that if the numerator in (E.13) is treated as a function of x, y and z, say $g(x,y,z)$, then volume_1 is the hypervolume under the hypersurface described by $g(x,y,z)$.

The *a posteriori* pdf for the general *a priori* pdf can also be found from the *a posteriori*

pdf that was obtained using the widely uniform *a priori* pdf. That is to say, the *a posteriori* pdf from the left side of (E.13) calculated with the widely uniform *a priori* pdf on the right side of (E.13) can be directly converted to the *a posteriori* for any general *a priori* pdf on the right side of (E.13). This is done multiplying the left side of (E.13) by the general *a priori* pdf and then dividing by the appropriate constant, say volume_2, to make the hyper-volume under the product equal to one. The resulting equation is

$$f^{(\mathsf{general})}_{x,y,z\left|\widehat{\mathbf{ToA}}\right.}(x,y,z\big|\boldsymbol{\tau})\mathrm{d}t = \frac{f^{(\mathsf{uniform})}_{\widehat{\mathbf{ToA}}\left|\right._{x,y,z}}(\boldsymbol{\tau}\big|x,y,z)f^{(\mathsf{general})}_{x,y,z}(x,y,z)\mathrm{d}t}{\mathsf{volume}_2}, \tag{E.14}$$

where

$$\mathsf{volume}_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathsf{f}^{(\mathsf{uniform})}_{\widehat{\mathbf{ToA}}\left|\right._{\mathsf{x,y,z}}}(\boldsymbol{\tau}\big|\mathsf{x,y,z})\mathsf{f}^{(\mathsf{general})}_{\mathsf{x,y,z}}(\mathsf{x,y,z})\mathrm{dxdydz}.$$

Now consider the construction of an approximate *a posteriori* pdf for an *a priori* pdf with any general shape. It was just demonstrated that the *a posteriori* pdf for a general *a priori* pdf can be constructed from the *a posteriori* pdf constructed from a widely uniform *a priori* pdf. Since we have a Gaussian approximation for the *a posteriori* pdf calculated with the widely uniform *a priori* pdf on the right side of (E.13), that Gaussian approximation can be used in place of the actual *a posteriori* pdf in updated to the *a posteriori* pdf for the general shape. Therefore, the approximate *a posteriori* pdf for any general *a priori* pdf is obtained by replacing the *a posteriori* pdf on the right side of (E.14) with its approximation.

There is no significant computational advantage in using the Gaussian approximation if volume_2 in (E.14) has to be computed with numerical integration. However, if the general *a priori* pdf is jointly Gaussian, then the numerical integration can be avoided if the Gaussian approximation for $f^{(\mathsf{uniform})}_{\widehat{\mathbf{ToA}}\left|\right._{x,y,z}}(\boldsymbol{\tau}\big|x,y,z)$ is used. If the general *a priori* density is Gaussian with a mean vector μ_0 and covariance matrix C_0 and if the Gaussian approximation for $f^{(\mathsf{uniform})}_{\widehat{\mathbf{ToA}}\left|\right._{x,y,z}}(\boldsymbol{\tau}\big|x,y,z)$, which obviously is Gaussian, has mean vector μ_1 and covariance matrix C_1, then the approximation to *a posteriori* pdf for the general *a priori* pdf, which in this

special case is Gaussian, is also Gaussian with the mean vector

$$\mu_2 = C_1(C_0 + C_1)^{-1}\mu_0 + C_0(C_0 + C_1)^{-1}\mu_1,$$

and covariance matrix

$$C_2 = C_1(C_0 + C_1)^{-1}C_0,$$

where

C_0 is the covariance matrix for x, y and z from the Gaussian *a priori* pdf (i.e., the general *a priori* pdf in Gaussian).

C_1 is the covariance matrix for x, y and z from the Gaussian approximation to the *a posteriori* pdf. The mathematics that produced the Gaussian approximation inherently assumes a widely uniform *a prirori* pdf.

C_2 is the covariance matrix for x, y and z from the approximate *a posteriori* pdf (which will be Gaussian) developed using the Gaussian approximation as well as a Gaussian *a priori* pdf.

μ_0 is the mean vector for x, y and z from the Gaussian *a priori* pdf (i.e., the general *a priori* pdf is Gaussian in this special case).

μ_1 is the mean vector for x, y and z from the Gaussian approximation to the *a posterirori* pdf. The mathematics that produced the Gaussian approximation inherently assumes a widely uniform *a prirori* pdf.

μ_2 is the mean vector for x, y and z from the approximate *a posteriori* pdf (which will be Gaussian) developed using the Gaussian approximation as well as a Gaussian *a priori* pdf.

References

[1] "SIGFOX." Available: `https://www.sigfox.com/`. [Online].

[2] Ingenu, "RPMA technology for the Internet of Things." Available: http://www.meterlinq.com/wp-content/uploads/2016/07/RPMA-Technology-Ingenu.pdf. [Online].

[3] M. Weyen, G. Ergeerts, R. Berkvens, B. Wojciechowski, and Y. Tabakov, "DASH7 alliance protocol 1.0: Low-power, mid-range sensor and actuator communication," in *IEEE Standards for Communications and Networking*, Oct. 2015.

[4] *White paper: A technical overview of LoRa and LoRaWAN*. LoRa Alliance, San Ramon, CA, USA, Tech. Rep., 2016.

[5] M. Chen, Y. Miao, Y. Hao, and K. Hwang, "Narrow band Internet of Things," *IEEE Access*, vol. 5, pp. 20557–20577, Sept. 2017.

[6] P. Sommer, Y. Maret, and D. Dzung, "Low-power wide-area networks for industrial sensing applications," *IEEE International Conference on Industrial Internet*, pp. 23–32, Oct. 2018.

[7] "GPS Accuracy." Available: `https://www.gps.gov/systems/gps/performance/accuracy/`. [Online].

[8] M. Lombardi, L. M Nelson, A. N Novick, and V. Zhang, "Time and frequency measurements using the global positioning system," *International Journal of Metrology*, vol. 8, pp. 26–33, July 2001.

[9] "Elecdata." Available: `http://elecdata.com/gps/technote/GPS_Rec_Cmp.pdf`. [Online].

[10] *LoRaWAN Geo-Location whitepaper*. LoRa Alliance Strategy Commitee, Jan. 2018.

[11] B. A. Renfro, M. Stein, E. B. Reed, J. Morales, and E. J. Villalba, *An Analysis of Global Positioning System (GPS) Standard Positioning Service Performance for 2019.* Space and Geophysics Laboratory, The University of Texas at Austin, May 2020.

[12] A. Coluccia and F. Ricciato, "A RSS-based localization via Bayesian ranging and iterative least squares positioning," *IEEE Commun. Letters*, vol. 18, pp. 873–876, May 2014.

[13] A. Coluccia, "RSS-based localization via Bayesian ranging and iterative least squares positioning," *IEEE Commun. Letters*, vol. 18, pp. 873–876, May 2014.

[14] J. Jiang, X. Y. Zheng, Y. F. Chen, C. H. Wang, P. T. Chen, C. L. Chuang, and C. P. Chen, "A distributed RSS-based localization using a dynamic circle expanding mechanism," *IEEE Sensors Journal*, vol. 13, pp. 3754–3766, Oct. 2013.

[15] Y. I. Wu, H. Wang, and X. Zheng, "WSN localization using RSS in three-dimensional space–A geometric method with closed-form solution," *IEEE Sensors Journal*, vol. 16, pp. 4397–4404, June 2016.

[16] L. Gui, M. Yang, and S. Yang, "RSS-based indoor localisation using MDCF," *IET Wireless Sensor Systems*, vol. 7, pp. 98–104, Apr. 2017.

[17] N. H. Pham, H. H. Nguyen, R. Barton, and J. Henry, "A novel ranging algorithm for RSS-based localization," *Workshop on Positioning, Navigation and Communication*, Oct. 2018.

[18] M. Compagnoni, A. Pini, A. Canclini, P. Bestagini, F. Antonacci, and S. Tubaro, "A geometrical-statistiscal approach to outlier removal for TDoA measurements," *IEEE Trans. Signal Process.*, vol. 65, pp. 3960–3975, Aug. 2017.

[19] A. A. Momtaz, F. Behnia, R. Amiri, and F. Marvasti, "NLoS identification in range-based source localization: Statistical approach," *IEEE Sensors Journal.*, vol. 18, pp. 3745–3751, May 2018.

[20] X. Li, X. Cai, Y. Hei, and R. Yuan, "NLoS identification and mitigation based on channel state information for indoor WiFi localisation," *IET Communications.*, vol. 11, pp. 531–537, Mar. 2017.

[21] W. Wang, G. Wang, F. Zhang, and Y. Li, "Second-order cone relaxation for TDoA-based localization under mixed LoS/NLoS conditions," *IEEE Signal Process. Letters*, vol. 23, pp. 1872–1876, Dec. 2016.

[22] A. N. Bishop, B. Fidan, B. D. O. Anderson, K. Dogancay, and P. N. Pathirana, "Optimal range-difference-based localization considering geometrical constraints," *IEEE Journal of Oceanic Engineering.*, vol. 33, pp. 289–301, July 2008.

[23] T. Qiao and H. Liu, "An improved method of moments estimator for ToA based localization," *IEEE Commun. Letters*, vol. 17, pp. 1321–1324, July 2013.

[24] Y. Zou, H. Liu, W. Xie, , and Q. Wan, "Semidefinite programming methods for alleviating sensor position error in TDoA localization," *IEEE Access*, vol. 5, pp. 23111–23120, Sept. 2017.

[25] L. Cong and W. Zhuang, "Non-line-of-sight error mitigation in TDoA mobile location," in *Proc. IEEE Global Telecommun. Conf.*, pp. 680–684, Nov. 2001.

[26] T. K. Le, K. Ho, and T. H. Le, "Rank properties for matrices constructed from time differences of arrival," *IEEE Trans. Signal Process.*, vol. 66, pp. 3491–3503, July 2018.

[27] A. Coluccia and A. Fascista, "On thehHybrid ToA/RSS range estimation in wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 17, pp. 361–371, Jan. 2018.

[28] A. Gholoobi and S. Stavrou, "Accelerating ToA/TDoA packet based localization methods," *IEEE Conference on Wireless Sensors*, Oct. 2014.

[29] W. Gerok, J. Peissig, and T. Kaiser, "TDOA assisted RSSD localization in UWB," *Workshop on Positioning, Navigation and Communication*, Mar. 2012.

[30] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. 4th Ed. McGraw-Hill, Inc., 2001.

[31] S. A. Zekavat and R. M. Buehrer, *Handbook of Position Location: Theory, Practice, and Advances*. 1st ed. John Wiley and Sons, Inc., 2012.

[32] R. Kaune, J. Horst, and W. Koch, "Accuracy analysis for TDoA localization in sensor networks," in *14th Int. Conf. on Information Fusion*, pp. 1647–1654, July 2011.

[33] J. Huang and Q. Wan, "Analayis of TDoA and TDoA/RSS-based geolocation techniques in a non-line-of-sight environment," *Journal of Communications and Networks.*, vol. 14, pp. 533–539, Oct. 2012.

[34] J. C. Chen, R. E. Hudson, and K. Yao, "Maximum-likelihood source localization and unknown sensor location estimation for wideband signals in the near field," *IEEE Trans. Signal Process.*, vol. 50, pp. 1843–1854, Aug. 2002.

[35] A. J. Fenwick, "Algorithms for position fixing using pulse arrival times," *IEE Proceedings - Radar, Sonar and Navigation.*, vol. 145, pp. 208–212, Aug. 1999.

[36] F. K. W. Chan, H. C. So, J. Zheng, and K. W. K. Lui, "Best linear unbiased estimator approach for time-of-arrival based localization," *IET Signal Process.*, vol. 2, pp. 156–162, July 2008.

[37] Y. T. Chan and K. C. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Process.*, vol. 42, pp. 1905–1915, Aug. 1994.

[38] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, "Least squares algorithms for time-of-arrival based mobile location," *IEEE Trans. Signal Process.*, vol. 52, no. 4, pp. 1121–1128, Apr. 2004.

[39] K. W. Cheung, H. C. So, W. K. Ma, and Y. T. Chan, "A constrained least squares approach to mobile positioning: algorithms and optimality," *EURASIP J. Adv. Signal Process.*, pp. 1–23, Dec. 2006.

[40] Q. Wan, Y. J. Luo, J. Xu, J. Tang, and Y. N. Peng, "Mobile localization method based on multidimensional similarity analysis," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing.*, pp. 1081–1084, Mar. 2005.

[41] H. C. So and F. K. W. Chan, "A generalized subspace approach for mobile positioning with time-of-arrival measurements," *IEEE Trans. Signal Process.*, vol. 55, pp. 5103–5107, Oct. 2007.

[42] K. W. Cheung and H. C. So, "A multidimensional scaling framework for mobile location using time-of-arrival measurements," *IEEE Trans. Signal Process.*, vol. 55, no. 2, pp. 460–470, 2005.

[43] H. W. Wei, Q. Wan, Z. X. Chen, and S. F. Ye, "Multidimensional scaling-based passive emitter localisation from range-difference measurements," *IET Signal Process.*, vol. 2, pp. 415–423, Jan. 2008.

[44] E. Xu, Z. Ding, and S. Dasgupta, "Source localization in wireless sensor networks from signal time-of-arrival measurements," *IEEE Trans. Signal Process.*, vol. 59, pp. 2887–2897, June 2011.

[45] L. Doherty, K. Pister, and L. E. Ghaoui, "Convex position estimation in wireless sensor networks," *IEEE INFOCOM*, vol. 3, pp. 1655–1663, Apr. 2001.

[46] P. Tseng, "Second-order cone programming relaxation of sensor network localization," *SIAM J. Optim.*, vol. 18, pp. 156–185, Aug. 2008.

[47] Z. Su, G. Shao, and H. Liu, "Semidefinite programming for NLoS error mitigation in TDoA localization," *IEEE Trans. Wireless Commun.*, vol. 22, pp. 1430–1433, July 2018.

[48] G. Wang, A. M. C. So, and Y. Li, "Robust convex approximation methods for TDoA-based localization under NLoS conditions," *IEEE Trans. Signal Process.*, vol. 64, pp. 3281–3296, July 2016.

[49] W. D. Foy, "Position-location solutions by Taylor-series estimation," *IEEE Trans. on Aerospace and Electronic Systems.*, vol. AES-12, pp. 187–194, Jan. 1976.

[50] R. Price, "A letter from the late Reverend Mr. Thomas Bayes, F. R.S to John Canton, M.A, and F.R. S," *Philosophical Transactions of the Royal Society of London*, pp. 269–271, Jan. 1763.

[51] H. L. V. Trees, *Detection Estimation and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory.* John Wiley & Sons, Inc., 1968.

[52] H. L. V. Trees, *Detection Estimation and Modulation Theory: Radar-Sonar Signal Processing and Gaussian Signals in Noise.* John Wiley & Sons, Inc., 2001.

[53] H. L. V. Trees, K. L. Bell, and Z. Tian, *Detection Estimation and Modulation Theory, Part I: Detection, Estimation, and Filtering Theory.* John Wiley & Sons, Inc., 2013.

[54] B. Mood, Graybill, and Boes, *Introduction to the Theory of Statistics.* McGraw-Hill, Inc., 1974.

[55] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice Hall, Inc., 1993.

[56] R. C. Shit, S. Sharma, D. Puthal, and A. Y. Zomaya, "Location of Things (LoT): A review and taxanomy of sensors localization in IoT infrastructure," *IEEE Communications Surveys and Tutorials.*, vol. 20, pp. 2028–2061, Third quarter 2018.

[57] C. Knapp and G. Carter, "The generalized correlation method for estimation of time delay," *IEEE Trans. Acous., Speech, Sig. Proccess.*, vol. 24, pp. 320–327, Aug. 1976.

[58] G. Carter, "Time delay estimation," *IEEE Trans. Acous., Speech, Sig. Proccess.*, vol. ASSP-29, pp. 461–462, June 1981.

[59] H. C. So, Y. T. Chan, and F. K. W. Chan, "Closed-form formulae for optimum time difference of arrival arrival based localization," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2614–2620, June 2008.

[60] S. R. Drake and K. Dogancay, "Geolocation by time difference of arrival using hyperbolic asymptotes," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing.*, pp. 361–364, May 2004.

[61] R. M. Vaghefi and R. M. Bueher, "Asynchronous time-of-arrival-based source localization," *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing.*, pp. 4086–4090, May 2013.

[62] M. Pauley and J. H. Manton, "The existence question for maximum-likelihood estimators in time-of-arrival-based localization," *IEEE Signal Process. Letters*, vol. 25, pp. 1354–1358, Sept. 2018.

[63] G. Wang, S. Cai, Y. Li, and M. Jin, "Second-order cone relaxation for TOA-based source localization with unknown start transmission time," *IEEE Trans. Veh. Technol.*, vol. 63, pp. 2973–2977, July 2014.

[64] G. Jacovitti and G. Scarano, "Discrete time techniques for time delay estimation," *IEEE Trans. Signal Process.*, vol. 41, pp. 525–533, Feb. 1993.