Efficient Techniques and Algorithms for Sensor Network Localization

Wenchao Li

ORCID: 0000-0001-8926-5539

Submitted in partial fulfilment of the requirements of the degree of

Doctor of Philosophy

Department of Electrical and Electronic Engineering
THE UNIVERSITY OF MELBOURNE

Jul 2018

Abstract

This thesis studies the phase ambiguity issues and practical solutions when radio waves are used to measure distance and time with applications to the localization problems in a wireless sensor network. The work involves theory extension and algorithm development.

Inter-sensor measurements, such as phase of arrival, time of arrival and angular of arrival from sensors with known locations to a free sensor node, are often used as an effective and convenient means for sensor node localization or related parameters in wireless sensor networks. Fundamentally, all of these measurements devolve to the signal , which is always wrapped to be within $[-\pi, \pi]$. In most practical applications, one need to unwrap the signal phases in order to calculate the distance between sensors. Signal with a wrapped phase is also known as phase ambiguity in literature. Resolving phase ambiguity problem requires to solve a congruence equation on integers and is traditionally addressed by the Chinese Remainder Theorem(CRT). However, the solution efficiency and robustness in practical algorithm development, in particular, when measurement noise is involved, become the major issues, and these have dominated the research discussion and direction in recent research literature.

In the thesis, the phase ambiguity problem is described with an application background of sensor localization in a wireless sensor network. The solutions, especially, the statistical solutions under the frameworks of both the Chines Remainder Theorem and lattice theory are discussed. The research starts from the investigation of a closed-form CRT algorithm and an modified CRT algorithm for resolving signal phase ambiguity in the presence of measurement noise is developed. It is analytically shown that the probability of success rate for integer reconstruction can be improved with slightly more

computational complexity. In the consequent study, research is done in the lattice theory approach, and it is shown that a stochastic phase ambiguity problem can be well and fits well into the framework of lattice theory. A new condition for the relaxation of the crucial condition in the early lattice algorithm is derived, which enables the development of the efficient lattice-based algorithm with more flexibility on the selection of parameters. Meanwhile, a fast iterative lattice algorithm using the concept of a Voronoi cell is proposed when the parameters are selected arbitrarily. It is also demonstrated that by taking into account the available variations of lattice structures and bases, an estimator with multiple candidate hypothesis represented by lattice points is proposed aiming at increasing the probability of successful reconstruction for different application situations. Several enhanced alternative algorithms, which are shown to be more efficient, are also developed.

In parallel with the solutions of key issues of signal phase ambiguity, the major techniques and algorithms for sensor localization in wireless sensor networks are presented. As an application example, a robot finding an object's location based on a wireless localization model, where localization noise depends on the distance between the robot and object, is formulated and analyzed in this thesis. This work demonstrates the importance of the analytical method, system modelling and statistical signal processing techniques presented in this thesis. In particular, in this thesis, two cases, static case and dynamical case, in the finding location problem are considered. It is shown that, by using statistical tools, the maximum likelihood estimator (MLE) is consistent in both cases under some mild conditions, which provide crucial information about the asymptomatic behaviour of the MLE under the complex situation.

Declaration

This is to certify that

1. the thesis comprises only my original work towards the PhD,
2. due acknowledgement has been made in the text to all other material used,
3. the thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Wenchao Li, Jul 2018

Acknowledgements

I would like to express my appreciation to all of the people who helped me to finish my PhD study and this thesis. I want to thank my two supervisors Prof. Bill Moran[1] and Prof. Timothy Brown[2], who spent lots of time on guiding my research and let me how to be a true researcher. It is indeed my great fortune to work with these two extraordinary mathematicians.

I would also like to express my special thanks to Dr. Xuezhi Wang[3] who not only helped me on research study, but also gave me a profusion of helpful suggestions on oversea life.

I would also like to thank the most important person around me, my wife Bihang Zhang. Thank you for supporting my work and taking care of our family to make me concentrate on my research.

My loving parents had always been a source of great strength, inspiration, and encouragement for me throughout my life. I am very blessed and fortunate to be with wonderful parents in law. Thanks for all the love and support you have given me over the years.

[1]Department of Electrical & Electronic Engineering, The University of Melbourne
[2]School of Mathematics and Statistics, The University of Melbourne
[3]School of Engineering, RMIT University

Preface

This thesis contains publications of which I am the primary author. The proportion of the contribution of me and co-authors in these publications is listed in following Table .

Title	Author	Contribution	Included in
"Efficient distance estimation using wrapped phase measurements," IET International Radar Conference 2015, Hangzhou, 2015, pp. 1-6. Oral presentation	Wenchao Li Xuezhi Wang Bill Moran	80% 10% 10%	Section 4.2, 4.3
"A lattice algorithm for optimal phase unwrapping in noise," 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016), Shanghai, 2016, pp. 2896-2900. Oral presentation	Wenchao Li Xuezhi Wang Bill Moran	80% 10% 10%	Section 4.3
"A lattice method for resolving range ambiguity in dual-frequency RFID tag localisation," 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2017), New Orleans, LA, 2017, pp. 3156-3160. Oral presentation	Wenchao Li Xuezhi Wang Bill Moran	80% 10% 10%	Section 4.4

Title	Author	Proportion of contribution	Included in
"Wireless signal travel distance estimation using non-coprime Wavelengths," IEEE Signal Processing Letters, vol. 24, no. 1, pp. 27-31, Jan. 2017.	Wenchao Li Xuezhi Wang Bill Moran	80% 10% 10%	Section 4.2
"A modified signal phase unwrapping algorithm for range estimation," 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), Calgary, Canada, 2018. Accepted.	Wenchao Li Xuezhi Wang Bill Moran	80% 10% 10%	Section 3.3
"An improved lattice algorithm and its performance analysis, " To be submitted to IEEE signal processing letter	Wenchao Li Xuezhi Wang Bill Moran	80% 10% 10%	Section 4.5

The second topic of this thesis (Part III), finding a position with distance dependent noise, is a joint work with Prof. Timothy Brown and Prof. Bill Moran, where Prof. Timothy Brown made a great contributions including developing the key ideas and providing constructive comments in proving the lemmas and theorems.

I am grateful of the grant support from Data61, former known as National ICT Australia(NICTA), and University of Melbourne.

Contents

List of Figures

Part I

Introduction

Chapter 1
Introduction

1.1 Background of Wireless Sensor Network

In the last few decades, due to the advances in processor technologies, wireless communications and accurate sensors that led to the development of small, low cost and power efficient sensor nodes, we have witnessed huge progress in Wireless Sensor Network (WSN) techniques. WSN has numerous applications in real life, from environmental monitoring, disaster monitoring, safety, security and smart grid to intelligent buildings, smart cars, human tracking and entertainment. The main applications are shown in Fig. 1.1.

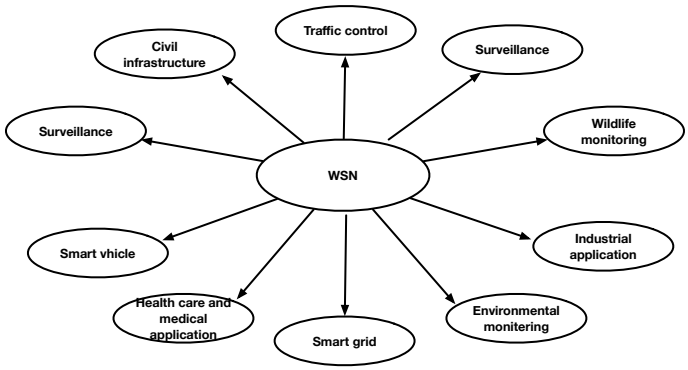


Figure 1.1: The possible applications of WSN

One of the main jobs of WSN is to collect physical data using various built-in sensors and transmit the information to central server for processing via wireless communication. Some WSN nodes has computing ability and can process the collected information locally. In the fields of WSN, the main research topics can be categorised under four

headings:

1. **Sensor localization**. In many applications of WSN, the sensor nodes are deployed randomly in a certain area or the sensor nodes are movable. Therefore, before using the sensed data provided by nodes, the positions of those nodes need to be determined, so as to enhance the value of the data. The Global Positioning System(GPS) is a possible method to obtain the sensor position, however in large scale networks or areas such as forests or indoors, the GPS is often not available because of higher cost or low signal strength. Thus, the ability to self-localize is highly desirable. One of the widely used strategies to achieve self-localization is to allow some sensors, called *anchors*, to have their position information at all times via GPS or manual deployment, and to localise other sensor nodes by the use of distance measurements, geometrical arguments and signal processing.

2. **Energy management**. In WSN, typically the sensor nodes are low cost and driven by batteries. When they are deployed in inaccessible regions, such as mountainous areas or hazardous environments, one of the main issues is the life of sensors which is largely dependent on the life of batteries. Therefore, in order to improve the performance of the WSN, the optimization of energy usage is of importance.

3. **Routing and communication**. Design of a priority routing protocol, aiming to maximize communication ability and battery life, is a crucial challenge in WSN.

4 **Sensor deployment**. Coverage and connectivity are the main requirements for the deployment of a WSN. To cover a maximum area using a limited number of sensors, the sensors should be widely spread, but this leads to communication problems.

1.2 Main Localization Techniques in Wireless Sensor Network

Among the research topics of WSN, sensor localization is crucial as it has great influence on the availability of sensor data. In general, the WSN localization algorithms estimate the locations of sensors with initially unknown location information by using positional knowledge of the anchors and inter-sensor measurements such as distance and bearings

information[2, 3]. Another similar scenario is to use sensor(or robot) to localize a static target. This case is very similar to the sensor localization using distance, and will be discussed later, if the unknown sensor is viewed as the static target and unable to communicate with anchors.

Generally speaking, the sensor localization techniques can be classified into two categories, i.e. range-based algorithm an range-free algorithm.

1. **Range-based algorithm**. In the range-based algorithm, the methodologies used to estimate the location of unknown sensors rely on the geometric properties of the sensors, such as triangulation and trilateration. The sensor measurements used to estimate the unknown location of sensors can be classified into two main types:

   1) Angle information or angle-of-arrival (AOA). Angular measurements are also referred to as bearings measurement. AOA is the angle between the propagation direction of an incident wave and some reference direction. The angle is presented in degree, (°), in clockwise direction and the 0° is generally defined as the North. Directional antennas or sensor arrays are usually used to obtain the measurement of AOA via calculating the phase difference between sensor array.

   2) Distance information. The distance measurement is defined as the relative distance between the anchor and unkonwn sensor which could be obtained via various techniques including phase measurements[4], one-way propagation time measurements, round-trip propagation time measurements[5, 6], time-difference of arrival (TDOA) and radio signal strength indicator (RSSI)[7], etc..

   As shown in Fig. 1.2, suppose that the unknown sensor is the emitter and the sensor 1, 2 and 3 are receivers, then angles θ_1, θ_2 and θ_3 are the measurements of AOA and the distance d_1, d_2 and d_3 are the measurements.

2. **Range-free algorithm**. In range-free localization, the positions of unknown sensors in the network are estimated using the network topology and the proximity relationship among nodes. The related algorithms include map-based algorithm,

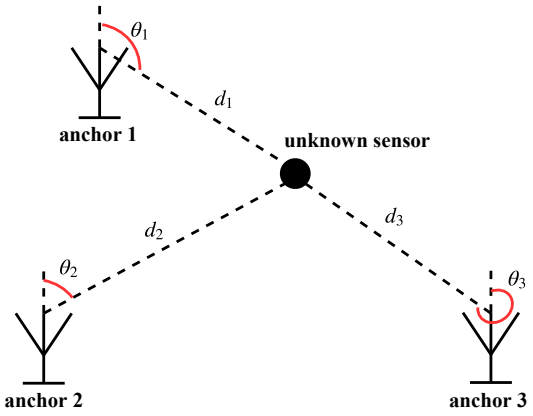


Figure 1.2: The angle and distance measurement of WSN

Distance Vector hop (DV-hop) localization, amorphous localization and Approximate Point in Triangulation (APIT). Range-free algorithms require less hardware, therefore, are more easily implemented and of lower cost. However, their accuracy is generally worse than range based localization. Range-free localization has attracted much research interest in recent decades [8–10].

The Fig.1.3 illustrates the main categorizations of self-localization in networks, along with examples of the forms of measurements combined with possible localization algorithms.

1.3 Problem Statement

1.3.1 The Phase Ambiguity Problem

In range-based algorithms, a key issue is how to obtain the distance information between the sensor and anchor. One of the widely used approaches is to measure the phase or phase difference[11, 12], of the arrived signal. This approach is easily implemented and has relatively high accuracy. Examples of localization schemes or algorithms using phase measurment include Radio Interferometric Positioning System (RIPS)[13], Radio Frequency IDentification (RFID)[14], mobile sensor localization[15] and acoustic localization[16].

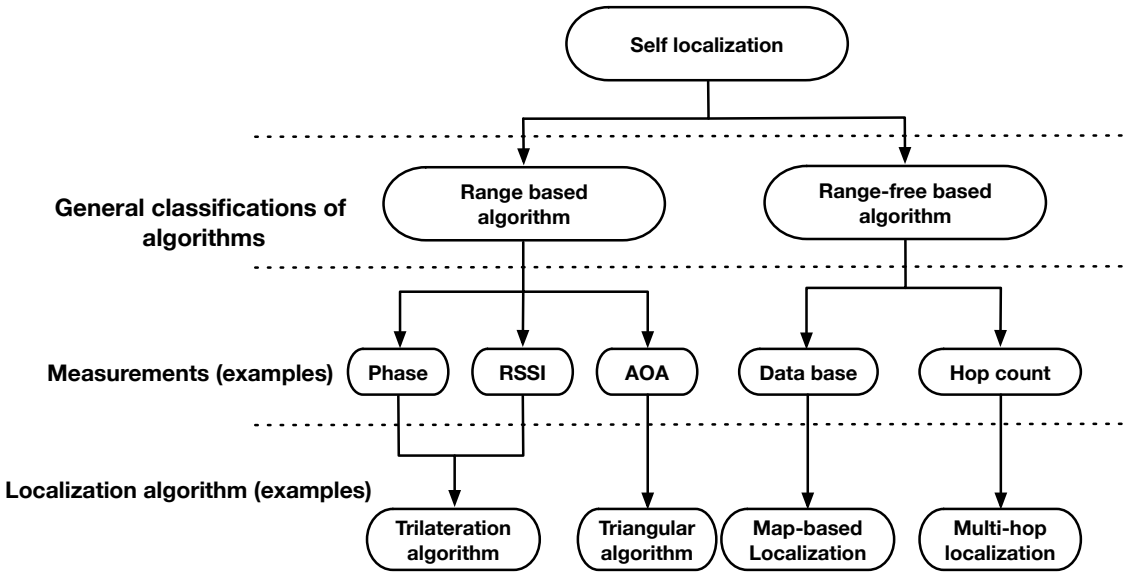


Figure 1.3: The taxonomy of sensor localization algorithm[1]

In general, phase measurements are obtained by measuring the arrived phase of the received signal or the difference between phases of two arriving signals. Then the relative distance can be calculated using the relationship between the phase and distance.

There are some issues in the phase-related localization algorithm, such as phase noise [17], time/frequency synchronization[18, 19], phase ambiguity[13, 15, 20, 21] and multi-path channel[22, 23]. Among these problems, phase ambiguity has the most influence on the localization accuracy since the phase measurement is limited the the range $[-\pi, \pi]$, i.e. wrapped by 2π, and then the true distance between the anchor and unknown sensor cannot be calculated directly. Fig.1.4 shows an example of phase ambiguity in sensor localization using a pure sinusoidal signal. In Fig.1.4, suppose that sensor **A** transmits a pure sine signal with wavelength λ and sensor **B** measures this signal in order to determine the distance r between them. However, sensor **B** can only measure the phase offset of this signal, i.e. ϕ, and does not know how many wavelengths are involved in this transmission. Therefore, the distance between **A** and **B** can not be determined exactly. This phenomenon can be formulated mathematically as follows:

$$r = n\lambda + \phi \tag{1.1}$$

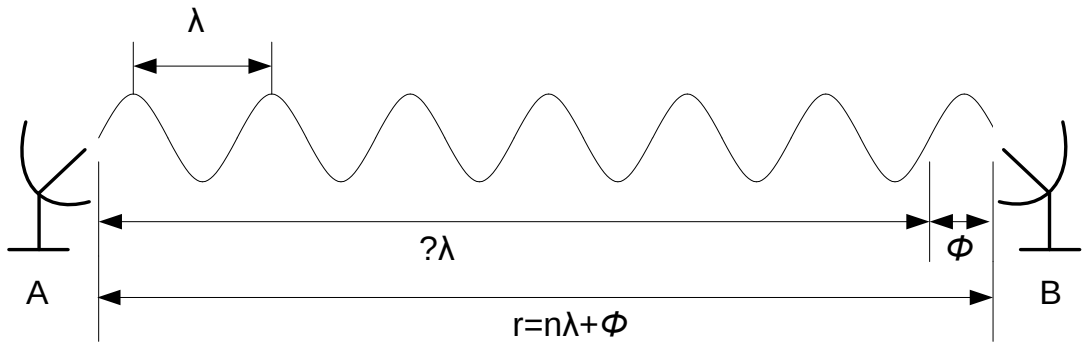


Figure 1.4: The example of phase ambiguity

where r is the true distance to be determined, $n \in \mathbb{Z}$ is an unknown number. It can be noticed that, in Eq.(1.1), different values of n will result in different values of distance with the same phase measurement ϕ. As a result, the phase measurement can not be used directly in localization algorithms because the true distance has numerous possible values. In consequence, efficient unwrapping or disambiguation techniques are required to ascertain the true distance before the localization algorithm can be implemented.

In order to solve the unknown distance r, or equivalently n, multiple phase measurements are obtained by using distinct wavelengths and therefore, a system of *congruence equations* can be constructed as

$$\begin{cases} r = n_1\lambda_1 + \phi_1 \\ \vdots \\ r = n_m\lambda_m + \phi_m \end{cases} \tag{1.2}$$

where $n_1, \cdots, n_m$ are unknown ambiguous integers, $\lambda_1, \cdots, \lambda_m$ are distinct wavelengths and $\phi_1, \cdots, \phi_m$ are phase measurements. A crucial concept in resolving the ambiguous r is the maximum unambiguous distance within which r can be ascertained uniquely from equation (1.2). This distance is the least common multiplier (LCM) of the wavelengths used:

$$r < \text{LCM}(\lambda_1, \cdots, \lambda_m) \tag{1.3}$$

In the ideal case, i.e the noise-free case, the well-known Chinese Remainder Theorem

(CRT) is good enough to resolve Eq.(1.2). However, in the real world, the phase measurements are contaminated by noise and, as a result, the conventional CRT is inapplicable since it is sensitive to noise[24] in the sense that a small error in measurement will lead to large deviation in estimation since the round operation is directly used in estimating the integer ϕ_i from noisy measurement.

Therefore, the algorithms which can efficiently estimate the distance from the noisy phase measurements are preferable. In general, there exist three frameworks to do so, i.e.

1. **Search-based algorithms**. The search-based algorithms compute all the possible values of unknown distance by searching over the space of the ambiguous integers. In general, the space of the ambiguous integers is huge and the searching process is time-consuming. The search-based algorithms are widely used in the case that the number of wavelengths used is small or the unknown distance is limited to a small region as it is easy to be implemented and understood.

2. **CRT-based algorithm**. The CRT-based algorithm is derived from modified CRT. It is shown to be more robust to noise than conventional CRT and has closed-form under mild conditions. The CRT-based algorithm is proved to have the same performance with the search-based. Some improvements to this algorithm were presented to solve ambiguity problem in different practical situations.

3. **Lattice-based algorithm**. Another candidate algorithm to solve the problem is based on lattice theory. In lattice theory, the objects of interest are discrete points arranged in a specific pattern in Euclidean space, and with surroundings that have identical features. The key idea of lattice-based algorithms in resolving ambiguity is to reformulate Eq.(1.2) as a closest vector (point) problem (CVP) in lattice theory. With some conditions, this CVP can be solved efficiently. The conventional lattice-based algorithm has the same computation complexity and performance with the CRT-based algorithm. However, because lattice structure can be easily visualized in low dimensions, and lattices have many good properties, lattice-based algorithms have a potential for adaptation to fit more situations.

One major focus of this thesis seeks to make contributions to resolve the ambiguous

distance from noisy phase measurements efficiently under different situations.

1.3.2 Find a Location When Noise Is Distance Dependent

The second major focus of this thesis is the theoretical analysis of static target localisation using a sensor on a moving platform. As mentioned above, localizing an target is akin to the localization of unknown sensor in a WSN, so the analysis of the former might help with understanding of the estimation problem.

Consider a specific scenario where a robot equipped with distance sensor along the real line and tries to track and reach the position of a fixed target, e.g. a door. An examples of position localization of robot in 1D case is shown in Fig. 1.5.

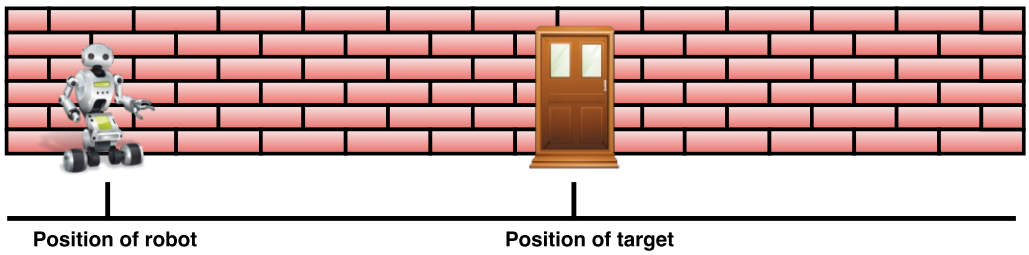


Figure 1.5: Illustration of position localization of robot

Suppose that the accuracy of the distance measurement is dependent on the relative distance between the robot and target, i.e. the closer the robot is to the target the better, the accuracy of the measurement. This noise model is found in many applications, in particular, in radar, sonar and lidar. A standard approach to estimation in these circumstances is to use Maximum Likelihood Estimation (MLE), or Maximum A Posteriori (MAP), the posterior mean, or some other standard statistical estimator. Assume that the robot moves to its estimated location of the target after each measurement and then performs another measurement. The question we address is whether this process converges, ultimately, so that the robot reaches the target. This seemingly simple problem is much more difficult than it seems. A proof of convergence in these circumstances is challenging. Mathematically this problem corresponds to the consistence of the estimator. A natural question is if the robot can grab the bean eventually when the Maximum

Likelihood Estimation is used? In mathematics, this question can be formulated as the consistency of the estimator.

Intuitively, there exist two strategies for robot to grab the bean, 1) keeps a static position and obtains the measurements until the estimation is consistent, and 2) updates robot's position using the estimation of bean at each step. We are interested in that, in either case, if the estimation is consistent eventually under some conditions.

1.4 Contributions

The major contributions of this thesis in resolving ambiguous distance falls into two categories: 1) relaxing the conditions on the conventional algorithms, and 2) improving the performance of algorithms in terms of increasing the successful reconstruction of the ambiguous integers and reducing computation complexity. In detail, the main contributions include:

1. **Relaxation of the coprimality condition in conventional lattice algorithm** [25, 26]

   One of the advantages of the conventional lattice algorithm proposed in [27] is that the algorithm returns the distance estimate using wrapped phases in $m-1$ steps, where m is the number of wavelengths used, with the same reconstruction performance of the search-based algorithm. However, in order to achieve this closed-form property, a crucial condition is that the wavelengths used have to be co-prime. In practice, this coprimality condition may prevent the algorithm being used as the wavelengths are in practice normally constrained to a limited bandwidth [13, 28], such as Wi-fi band (2.4GHz and 5GHz), or Industrial, Scientific and Medical (ISM) radio bands. In order to solve the distance uniquely, the unknown distance r has to be less than maximum unambiguous distance as shown in (1.3), thus the constrained bandwidth imposes limitations on the maximum unambiguous distance. For example, the admissible wavelengths in the 2.4GHz Wi-fi band are $\{125, 124, \cdots, 120\}$mm and thus the coprime wavelengths in this interval, e.g. $\{121, 120\}$mm, may not cover a sufficient unambiguous distance. In addition, the coprimality condition may reduce the resolution of range; for example, in a

medium-to-high Pulse repetition frequency (PRF) radar, the target may be unobservable because of the so-called "blind zone" and clutter [29] when the number of available PRFs is not enough. Clearly, the usefulness of the algorithm will be improved by the relaxation of the coprimality condition. Base on this, a relaxed and more general condition for the lattice algorithm is presented while retaining its closed form. This new condition allows a larger number of integer-related wavelengths to be used.

2. **A computationally efficient lattice-based algorithm with non-coprime wavelengths** [30]

   Under the relaxed condition mentioned above, the closed-form lattice algorithm can be used to resolve the ambiguous distance. However, in practice, the available wavelengths can not always satisfy this relaxed condition. One possible solution to this is the search-based algorithm. On the other hand, implementation requires high computational complexity because of the need for exhaustive searches over the parameter space. In this thesis, a general iterative lattice algorithm to solve this general ambiguity problem with the non-coprime walvengths is presented. Based on the work in [31], a modified iteration for closest point searching is derived so that the proposed algorithm is computationally more efficient than existing approaches.

3. **An improved lattice algorithm using different structures of lattice basis** [32]

   In the problem of resolving the ambiguous distance in localization, one major concern is the probability of successful reconstruction of wrapped integers in high noise levels, since if the unknown integers are estimated incorrectly, the distance estimation error can be very significant. In order to increase the probability of successful reconstruction, an improved lattice algorithm with slightly more computations is presented. In the lattice algorithm, the basis of the lattice space, which influences the performance of the algorithm in a noise circumstance is determined by the order of wavelengths used. Based on this, in the improved algorithm, the different structures of the lattice basis are obtained by the ordering of the wavelengths used and then more information about the wrapped integers are collected.

As a result, a set of candidate estimates of the integers are obtained according to the structures of lattice basis, then a criterion based on the noise distribution is presented to select a reasonable estimation from the candidates.

4. **An improved CRT algorithm** [33]

   As mentioned earlier, the probability of successful reconstruction of wrapped integers is of importance in localisation. Our contribution is to improve the performance of CRT in terms of the probability of successful reconstruction. In the existing CRT algorithm, the difference of two integral measurements is used to solve the wrapped integers. However, in practice, measurements are unnecessary to be integers due to the noisy, therefore, the difference of measurements are normally non-integer. The rounding operation applied to the fractional difference yields an estimate of the value of the integer. This estimate has a significant influence on the performance of CRT algorithm. Unfortunately, this rounding operation can estimate the difference erroneously in the high noise regime. By considering this, we propose a modified estimator which is able to compensate for rounding error and improve the estimate of the difference, thereby improving the performance of the CRT.

5. **A modified lattice algorithm by considering multiple candidate estimations** [26, 34]

   This contribution is to improve the performance of the lattice in terms of the probability of successful reconstruction. In contradistinction to the contribution described above, which operates by changing the lattice basis, in this case the basis is fixed, and multiple candidates estimations are obtained by using properties of the noise. The final estimation is selected using an optimization process. In the conventional lattice algorithm, a lattice point in a certain lattice basis is found which is the closest one to the given point calculated from the noisy measurements. The correct result is returned if the given point lies in the Voronoi cell of the lattice point which corresponds to the ground truth. However, due to the noise, the given point will not always lie in the correct cell, but in the neighbor ones. By considering this, we

could consider the neighbor lattice points of the given points as the candidates and then obtain the estimation using the noise property.

The second major focus of this thesis is an investigation of the behaviour of the MLE in finding the location of target using distance dependent noise:

6 **Finding the position when the noise is distance dependent**

We restrict attention to the one dimensional case, since even there the problem is surprisingly challenging. Consider the case of a robot moving a long the line and taking measurements of its distance to a fixed target. The robot measurement is the location of the target, relative to its position, which is unknown. The noise assumed Gaussian with mean 0 and the variance dependent on the distance between the robot and target. Conventionally, if normal distribution is considered, the MLE is consistent if the noise variance is constant. However, when the noise variance is dependent on the distance, the proof of consistency is, as far as we are aware, absent from the literature . In the Part III, we prove that the MLE, when the noise variance is dependent on the distance, is consistent if the robot is static or moving under some conditions via analyzing the likelihood function and the use of martingale theory.

Part II

Efficient Algorithms to Resolve Ambiguity

Chapter 2
Preliminaries

In this chapter, the necessary background of lattice theory including the basic definition, closest vector problem, shortest vector problem and some related algorithms, such as sphere decoder, Babai's algorithm and LLL algorithm, are introduced in Section 2.1. In Section 2.2, the localization example in wireless sensor network using phase difference measurement is introduced and the ambiguity problem is formulated. The classical Chinese Remainder Theorem is given in Section 2.3

2.1 Introduction to Lattice Theory

Lattice theory is a key part of number theory and studies the arrangement of points in Euclidean space. It has many applications in mathematics and engineering fields, e.g. the solution of integer programming problems[35], diophantine approximation[36, 37], cryptography[38, 39], telecommunication, the design of error correcting codes for multi antenna systems[40–43] and frequency estimation[44]. This section will cover the definition, basic problems and related algorithms of lattice.

2.1.1 The Definition of Lattice

Definition 2.1. A *lattice* $\mathbf{\Omega}$ is a discrete additive subgroup of $\mathbb{R}^m$, i.e., it is a subset $\mathbf{\Omega} \subset \mathbb{R}^m$ satisfying:

1. (subgroup) $\mathbf{\Omega}$ is closed under addition and subtraction operations;
2. (discrete) $\exists \epsilon > 0$ such that any two distinct lattice points $\boldsymbol{x} \neq \boldsymbol{y}$ and $\boldsymbol{x}, \boldsymbol{y} \in \mathbf{\Omega}$ satisfying

$$\|\boldsymbol{x} - \boldsymbol{y}\| \geq \epsilon$$

□

Remark 2.1. Not every subgroup of $\mathbb{R}^m$ is a lattice. □

Another definition is given by

Definition 2.2. Let $\mathbf{a}_1, \cdots, \mathbf{a}_n$ be be linear independent row vectors in $\mathbb{R}^m$ and $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$. The *lattice* $\mathbf{\Omega}$ is defined by

$$\mathbf{\Omega} \triangleq \{\boldsymbol{x} = \boldsymbol{n}\mathbf{A} | \boldsymbol{n} \in \mathbb{Z}^n\} \tag{2.1}$$

where $\mathbf{A}$ called as the *basis* or *generator* of lattice and the lattice generated by $\mathbf{A}$ is denoted by $\mathbf{\Omega}(\mathbf{A})$. □

For example, consider following bases of two lattices:

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } \mathbf{A}_2 = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \tag{2.2}$$

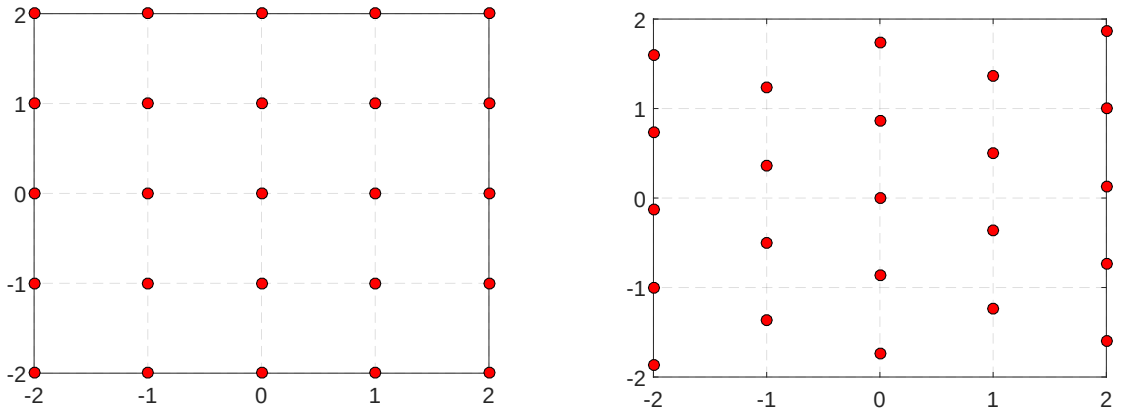
the lattice $\mathbf{\Omega}(\mathbf{A}_1)$ and $\mathbf{\Omega}(\mathbf{A}_2)$ are shown in Fig.2.1a and 2.1b, where the red dots are the lattice points generated by $\mathbf{A}_1$ and $\mathbf{A}_2$ respectively according to different value of $\boldsymbol{n} \in \mathbb{Z}^2$.

There are two fundamental problems in lattice theory, that concern us. The first is the shortest vector problem (SVP) and the second is the closest vector problem (CVP), also called closest point searching. In lattice theory, the SVP and CVP are central algorithmic problems which have applications in integer programming [45], factoring polynomials over the rationals [46], Cryptography [47–49], and much more.

2.1.2 Shortest Vector Problem

In a lattice, the basis is not unique: one could find many bases which generate the same lattice space. This enables us to change a ”bad” basis to a ”good” one.

There are standard elementary operations on a basis that will retain the basis property for a given lattice. They are:



(a) The lattice space with basis $\mathbf{A}_1$

(b) The lattice space with basis $\mathbf{A}_2$

Figure 2.1: Two examples of 2-dimensional lattice, the lattice points are denoted by red dots

1. interchanging two rows in $\mathbf{A}$;
2. multiplication of a row by -1;
3. multiplying a row by an integer and adding it to a second row, thereby replacing the second row.

Moreover, Lemma 2.1 is usually used in practice to find the alternative basis.

Definition 2.3. A unimodular matrix $\mathbf{U}$ is a square integer matrix having determinant ± 1; that is, $|\det(\mathbf{U})| = 1$. □

Lemma 2.1. Bases $\mathbf{A}$ and $\mathbf{B}$ generate identical lattice, i.e. $\mathbf{\Omega}(\mathbf{A}) = \mathbf{\Omega}(\mathbf{B})$ if and only if there exists a unimodular matrix $\mathbf{U}$ satisfying $\mathbf{B} = \mathbf{UA}$. □

A crucial application of CVP is known as *reduction* which aims to find a basis in which, informally, the basis consists of short vectors that are (almost) orthogonal to each other, whereas "bad" means that it consists of long vectors that generally point in the same (or opposite) direction. As so (informally) defined it is hard to find basis that is absolutely the best by this criterion.

Fig. 2.2 shows a reduction example. In the figure, the green vectors $[a'_1, a'_2]^T$ and the alternative blue vectors $[a_1, a_2]^T$ are both bases of the lattice and the blue ones are

evidently better than the green ones as they are shorter.

The *orthogonality defect* is a measure of the orthogonality of the basis.

Definition 2.4 (orthogonality defect)**.** The orthogonality defect of a basis $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 \\ \vdots \\ \mathbf{a}_n \end{bmatrix}$, $\gamma(\mathbf{A})$, is the quantity

$$\gamma(\mathbf{A}) = \frac{\prod_{i=1}^{m} \|\mathbf{a}_i\|}{|\det(\mathbf{A}^T\mathbf{A})} \tag{2.3}$$

□

Remark 2.2. It can be shown that $\gamma(\mathbf{A}) \geq 1$ and the equality holds if and only if the basis is orthogonal. □

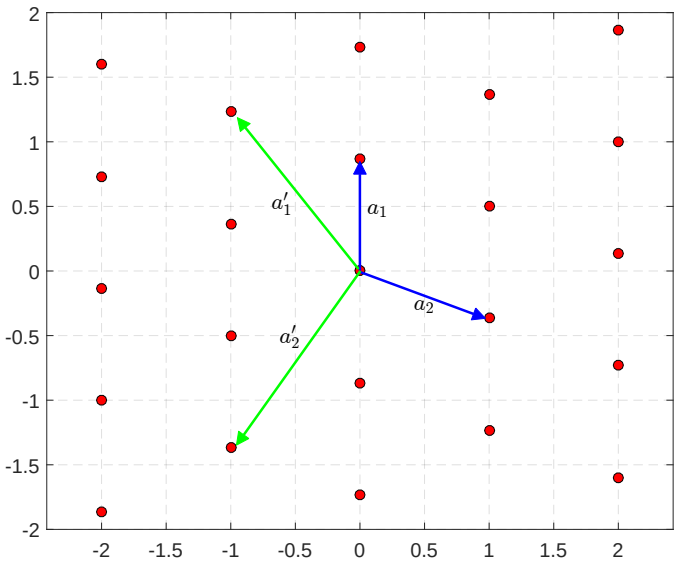


Figure 2.2: The example of basis reduction.

There exist many notions for reduction, such as Minkowski-reduced[50], Hermite-Korkine-Zolotarev (HKZ)-reduced[51] and Lenstra-Lenstra-Lovasz (LLL)-reduced bases[46]. Among them, the Minkowski reduction is the most intuitive one. However, as the dimension increases, the computation of Minkowski-reduced and HKZ-reduced basis is NP-Hard.

1. **Minkowski reduction.** Before introducing Minkowski reduction, some definitions are needed.

   Definition 2.5 (Successive Minima[52]). For a lattice $\mathbf{\Omega} \in \mathbb{R}^m$, let $\mathcal{B}(c,r)$ be a closed ball center at c with radius r, then define L_i as

   $$L_i = \{r : \mathcal{B}(0,r) \text{ contains at least } i \text{ linearly indep. lattice vectors}\} \tag{2.4}$$

   Then $L_1 \leq L_2 \leq \cdots \leq L_m$ are the successive minima of $\mathbf{\Omega}$ □

   Remark 2.3. L_1 is the Euclidean length of a shortest nonzero lattice vector of the given lattice $\mathbf{\Omega}$. Fig.2.3 shows an example of successive minima with basis $\left[1, \frac{1}{2}; 0, \frac{\sqrt{3}}{2}\right]$, and $L_1 = \frac{\sqrt{3}}{2}$, $L_2 = \sqrt{1 + \left(1/2 - \sqrt{3}/2\right)^2} \approx 1.065$. □

   Definition 2.6 (Minkowski-reduced basis[50]). A lattice basis $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m]$ is called Minkowski-reduced if, $\forall i$, $1 \leq i \leq m$, the vector $\mathbf{a}_i$ has the minimum norm among all lattice vectors $\mathbf{A}_i$ such that $[\mathbf{a}_1, \cdots, \mathbf{a}_i]$ can be extended to a basis for $\mathbf{\Omega}(\mathbf{A})$. □

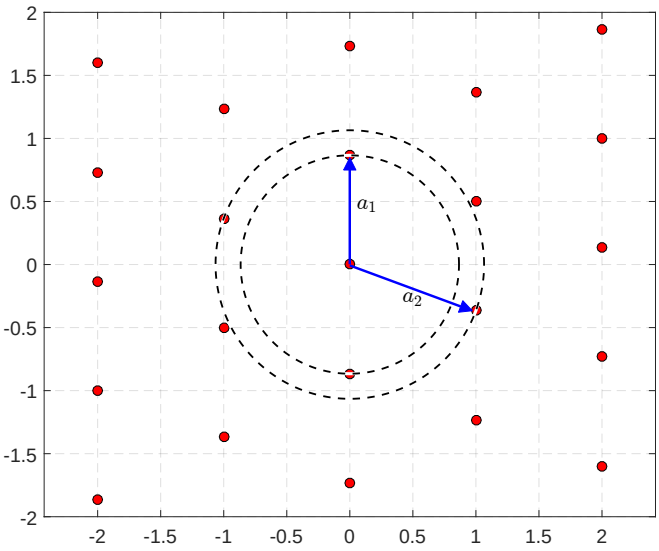


Figure 2.3: The example of successive minima with basis $\left[1, \frac{1}{2}; 0, \frac{\sqrt{3}}{2}\right]$.

The basis of an m-dimensional lattice that reaches the m successive minima must be Minkowski-reduced, however, a Minkowski-reduced basis may not reach suc-

cessive minima[53]. The Minkowski-reduced basis is optimal up to 4 dimension compared to other reduction algorithm[53,54].

Another result states that the orthogonality defect of a Minkowski-reduced basis can be upper-bounded by a constant that only depends on the lattice dimension.

2. **HKZ reduction.** Korkine and Zolotarev proposed a reduced basis by strengthening Hermites size-reduction[51, 55, 56]. In higher dimensions, the reduction of HKZ reduced basis is stronger than the Minkowski reduced basis in the sense that all the elements of a HKZ reduced basis are known to be very close to the successive minima, but the former one is expensive to compute. Next, we introduce the HKZ reduction and related definitions.

Definition 2.7 (Gram-Schmidt orthogonalization)**.** Let $[\mathbf{a}_1, \cdots, \mathbf{a}_m]$ be linearly independent vectors. The Gram-Schmidt orthogonalization $[\mathbf{a}_1^*, \cdots, \mathbf{a}_m^*]$ is: $\forall 1 \leq i \leq m$, $\mathbf{a}_i^*$ is the component of $\mathbf{a}_i$ that is orthogonal to the subspace spanned by the vector $[\mathbf{a}_1^*, \cdots, \mathbf{a}_{i-1}^*]$. Moreover, the Gram-Schmidt coefficients $\mu_{i,j}$ satisfy, $\forall 1 \leq i \leq m$

$$\mathbf{a}_i = \mathbf{a}_i^* + \sum_{j<i} \mu_{i,j} \mathbf{a}_j^* \tag{2.5}$$

and

$$\mu_{i,j} = \frac{\langle \mathbf{a}_i, \mathbf{a}_j^* \rangle}{\|\mathbf{a}_j^*\|} \quad \forall\, j < i \tag{2.6}$$

□

Definition 2.8 (QR decomposition)**.** For any $m \times n$ matrix $\mathbf{A}$ with $\text{rank}(\mathbf{A}) = n \leq m$, we can construct an $m \times n$ orthogonal matrix $\mathbf{Q}$ satisfying $\mathbf{Q}^T\mathbf{Q} = \boldsymbol{I}$, where $\mathbf{I}$ is identity matrix, and an $n \times n$ upper triangular matrix $\mathbf{R}$ such that $\mathbf{A} = \mathbf{QR}$. □

Then we have following definition:

Definition 2.9 (Size-reduced)**.** A lattice basis $\mathbf{A} \in \mathbb{R}^{m \times n}$ is called size-reduced if the upper triangular factor $\mathbf{R} = \{r_{i,j}\}$ of the QR decomposition, where $i = 1, \cdots, m$,

$j = 1, \cdots, n$ and $r_{i,j}$ are entries of $\mathbf{R}$, satisfies

$$|r_{i,j}| \leq \frac{1}{2} r_{i,i} \text{ for } 1 \leq i < j \leq n \tag{2.7}$$

□

Now, the definition of HKZ-reduced basis is as follows:

Definition 2.10 (HKZ reduction)**.** A lattice basis $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m]$ is said to be HKZ-reduced if

1) $\|\mathbf{a}_1\| = L_1(\mathbf{\Omega}(\mathbf{A}))$;
2) $\mathbf{a}_i$ is a lattice vector having minimal non-zero distance to the linear span of $(\mathbf{a}_1, \cdots, \mathbf{a}_{i-1})$ for $i \geq 2$;
3) the basis $\mathbf{A}$ is size-reduced.

□

3. **LLL reduction.** The LLL algorithm of Lenstra, Lenstra Lovász was presented in [46]. It provides reduced basis in polynomial time with proven quality. The LLL algorithm is based on the Lagrange-Gauss algorithm and Gram-Schmidt orthogonalization. Since there is no lattice basis reduction algorithm able to give optimal results in polynomial time, LLL algorithm, which provides an approximation of the shortest vector, is widely used in many applications.

   Definition 2.11 (LLL reduction)**.** Given a $\delta \in \left(\frac{1}{4}, 1\right)$ and a lattice basis $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m]$. Let the QR decomposition of $\mathbf{A}$ be $\mathbf{A} = \mathbf{QR}$ (Definition 2.8). Then $\mathbf{A}$ is called LLL-reduced if

   1. (size reduction) $\mathbf{R}$ is sized-reduced and
   2. (Lovász condition) $r_{i,i}^2 + r_{i-1,i}^2 \geq \delta r_{i-1,i-1}^2$

□

Remark 2.4. It is usual to choose $\delta = \frac{3}{4}$ in Lovász condition. It should be noticed that though the determinant of $\mathbf{Q}$ in QR decomposition of $\mathbf{A}$ is either 1 or -1, $\mathbf{R}$

cannot be used for lattice reduction as $\mathbf{Q}$ is generally not an integer matrix. □

The algorithm to compute the LLL-reduced basis given an original basis is provided in Algorithm 2.1 at the end of this chapter. The LLL algorithm is an iterative algorithm. As seen in the Definition 2.11, the algorithm falls into two main procedures: 1. the size reduction procedures and; 2. swap procedures (Lovász condition). Size reduction can be obtained by performing integer linear combinations. On the other hand, two columns, $\mathbf{a}_i$ and $\mathbf{a}_{i-1}$, in a basis are swapped if $\mathbf{a}_i$ is not not significantly longer than $\mathbf{a}_{i-1}$.

2.1.3 Closest Vector Problem

The Closest Vector Problem (CVP) is also known as the the nearest point searching problem. Given an m-dimensional lattice $\boldsymbol{\Omega}$ and an arbitrary point $\boldsymbol{x} \in \mathbb{R}^m$, the CVP aims to find a lattice point $\boldsymbol{p} \in \boldsymbol{\Omega}$ which is closest to $\boldsymbol{x}$ in the sense of Euclidean distance, i.e.

$$\boldsymbol{p} = \arg\min_{\boldsymbol{v}\in\boldsymbol{\Omega}}\{\|\boldsymbol{x}-\boldsymbol{v}\|\} \quad \text{for} \quad \boldsymbol{x} \in \mathbb{R}^m \tag{2.8}$$

CVP can also be described using Voronoi cells[58].

Definition 2.12 (Voronoi cell of lattice)**.** Given a lattice $\boldsymbol{\Omega} \subseteq \mathbb{R}^m$ and a point $\boldsymbol{p} \in \boldsymbol{\Omega}$, the Voronoi cell of $\boldsymbol{p}$, $\boldsymbol{V}(\boldsymbol{\Omega}, \boldsymbol{p})$, is defined by

$$\boldsymbol{V}(\boldsymbol{\Omega}, \boldsymbol{p}) = \left\{\boldsymbol{x} \in \mathbb{R}^m : \forall \boldsymbol{p}' \in \boldsymbol{\Omega}, \boldsymbol{p}' \neq \boldsymbol{p}, \|x-\boldsymbol{p}\| \leq \|x-\boldsymbol{p}'\|\right\} \tag{2.9}$$

□

Then the CVP is to find a Voronoi cell of the lattice point which the given point $x \in \mathbb{R}$ lies in. A solution of CVP is given in Fig. 2.4, where the shaded Voronoi cell is the solution of CVP.

Since Voronoi cell for any lattice points will be translates of the Voronoi cell of the origin, Voronoi cells for the lattice points will be translates of the Voronoi

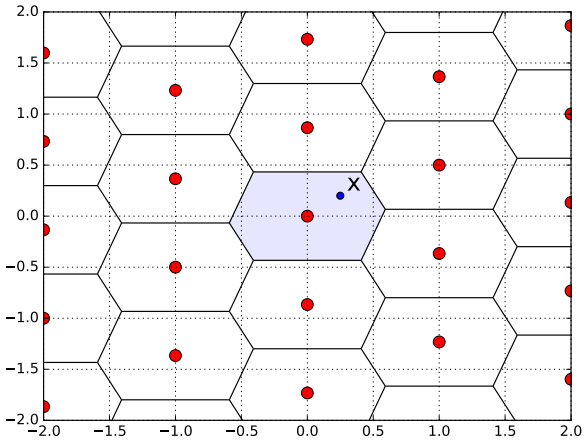


Figure 2.4: The example of CVP where the arbitrary point $x \in \mathbb{R}$ lies in the shaded Voronoi cell which is associated to a lattice point that is closest to the x

it suffices to calculate the latter. Thus we defined $\boldsymbol{V}(\boldsymbol{\Omega}) \triangleq \boldsymbol{V}(\boldsymbol{\Omega}, \mathbf{0})$. Another property related to Voronoi cell is the relevant vector.

Definition 2.13 (Relevant vector)**.** Given a lattice space $\boldsymbol{\Omega}$, the relevant vectors are those points $\boldsymbol{v} \in \boldsymbol{\Omega} \setminus \{\mathbf{0}\}$ such that

$$\boldsymbol{v} \cdot \boldsymbol{p} < \boldsymbol{p} \cdot \boldsymbol{p} \quad \forall \boldsymbol{p} \in \boldsymbol{\Omega} \setminus \{\mathbf{0}\} \tag{2.10}$$

□

An example of Voronoi cell and relevant vector is shown in Fig.2.5 where the red dots represent lattice points, the cell surrounded by the blue line segment is a Voronoi cell, and the black arrows are relevant vectors.

Then the CVP aims to find a Voronoi cell for an arbitrary point $\boldsymbol{x} \in \mathbb{R}^m$ such that

$$\boldsymbol{p} = \{\boldsymbol{v} | \boldsymbol{x} \in \boldsymbol{V}(\boldsymbol{\Omega}, \boldsymbol{v}), \forall \boldsymbol{v} \in \boldsymbol{\Omega}\} \quad \text{for} \quad \boldsymbol{x} \in \mathbb{R}^m \tag{2.11}$$

There exist many algorithms to solve the CVP, such as Babai's nearest-plane algorithm[59], the sphere decoder[60–62], and the slicing algorithm[63].

If the lattice basis chosen is a diagonal, then the result of Babai's algorithm is optimal,

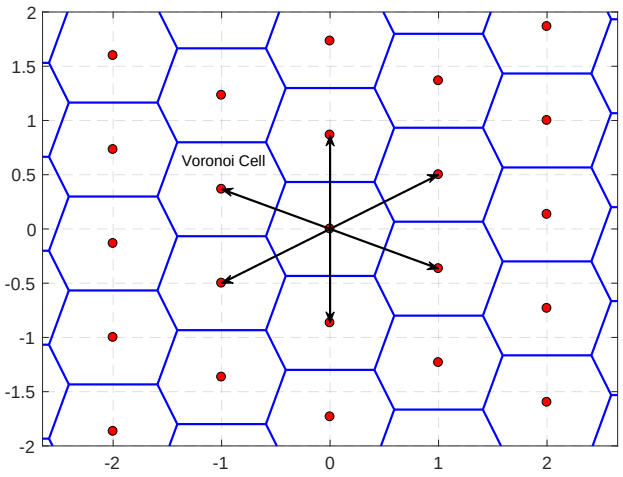


Figure 2.5: The example of Voronoi cell and relevant vector.

i.e. the output lattice point is the closest one to the given point. However, if the basis is non-diagonal, the output of Babai's algorithm is not guaranteed to be optimal. Babai's algorithm is described in Algorithm 2.2.

Algorithm 2.2: Babai's nearest-plane algorithm

Data: A lattice basis $\mathbf{A} = [\mathbf{a}_1, \cdots, \mathbf{a}_m]$ and given point $\boldsymbol{x} \in \mathbb{R}^m$

Result: closest lattice point $\boldsymbol{v}$

Compute Gram-Schmidt basis $[\mathbf{a}_1^*, \cdots, \mathbf{a}_m^*]$

$\boldsymbol{x}_m = \boldsymbol{x}$

for $i = m : -1 : 1$ **do**

  $l_i = \frac{\boldsymbol{x}_i \cdot \mathbf{a}_i^*}{\mathbf{a}_i^* \cdot \mathbf{a}_i^*}$

  $\boldsymbol{y}_i = \lfloor l_i \rceil \mathbf{a}_i$

  $\boldsymbol{x}_{i-1} = \boldsymbol{x}_i - (l_i - \lfloor l_i \rceil)\mathbf{a}_i^* - \lfloor l_i \rceil \mathbf{a}_i$

$\boldsymbol{v} = \sum_{i=1}^{m} \boldsymbol{y}_i$

where $\lfloor \cdot \rceil$ is rounding operation.

Another efficient algorithm is the sphere decoder and this is described in Algorithm 2.3 at the end of this chapter. In this case, $\mathbf{R}$ is the upper-triangular matrix obtaining from QR decomposition (see Definition 2.8) of the reduced lattice basis $\mathbf{A}$ and the closest point can be calculated from $\boldsymbol{v}\mathbf{A}$, where $\boldsymbol{v}$ is output of the algorithm. In this algorithm, the

function $sgn(\cdot)$ is used and defined by

$$sgn(x) = \begin{cases} -1 & x \leq 0 \\ 1 & x > 0 \end{cases}$$

2.2 The Ambiguity Problem in Sensor Localization

In a wireless sensor network (WSN), one of the main aims is to collect data from the environment such as temperature, air pressure, humidity etc.. However, before uploading these data to process centrally, it is important that the locations where the sensors collect these data should be identified. The sensors can be localized via Global Positioning System (GPS), although in a large sensor network with a high density of nodes deployment of many GPS can become expensive. There are also many circumstances in which GPS may be unavailable. A practical approach is to use a few sensors with known position, normally called anchors. The position might be obtained by onboard GPS or by manual deployment. The other nodes are then located by means of distance measurements to the other anchors. One such approach is the Radio Interferometric Positioning System (RIPS). This method has low cost and relatively high accuracy.

In this section, the RIPS will be used as an example to introduce the ambiguity problem in sensor localization, and then the rigorous mathematical definition of ambiguity is given.

2.2.1 The Ambiguity Problem in Sensor Localization

The concept of RIPS is firstly presented in [13] and extended in papers [64–68]. The localization technique is investigated in [69]. This technique requires at least 3 anchors to localize one unknown sensor node.

As shown in Fig. 2.6, sensors **A**, **B** and **C** are anchors and **D** is the sensor to be localized. Anchor **A** and **B** transmit pure sine signal with frequencies $f_{\mathbf{A}}$ and $f_{\mathbf{B}}$, then the

transmitted signal $s_{\mathbf{A}}(t)$ and $s_{\mathbf{B}}(t)$ can be represented as

$$s_{\mathbf{X}}(t) = a_{\mathbf{X}} \cos(2\pi f_{\mathbf{X}} t)$$

where $\mathbf{X} = \mathbf{A}$ or $\mathbf{B}$, $a_{\mathbf{X}}$ is the amplitude of the signal. These two frequencies satisfy

$$|f_{\mathbf{A}} - f_{\mathbf{B}}| \leq 1\text{kHz}$$

Then, anchor $\mathbf{C}$ and unknown sensor $\mathbf{D}$ act as receivers and the received signal, $s_{\mathbf{C}}(t)$ and $s_{\mathbf{D}}(t)$, can be written as follows in the absence of noise

$$s_{\mathbf{C}}(t) = a_{\mathbf{A}} \cos(2\pi f_{\mathbf{A}} t + \phi_{\mathbf{AC}}) + a_{\mathbf{B}} \cos(2\pi f_{\mathbf{B}} t + \phi_{\mathbf{BC}})$$

and

$$s_{\mathbf{D}}(t) = a_{\mathbf{A}} \cos(2\pi f_{\mathbf{A}} t + \phi_{\mathbf{AD}}) + a_{\mathbf{B}} \cos(2\pi f_{\mathbf{B}} t + \phi_{\mathbf{BD}})$$

where the phase offset $\phi_{\mathbf{XY}}$, $\mathbf{X} = \mathbf{A}$ or $\mathbf{B}$ and $\mathbf{Y} = \mathbf{C}$ or $\mathbf{D}$, due to the traveled distance is defined by

$$\phi_{\mathbf{XY}} = 2\pi \frac{f_{\mathbf{X}} d_{\mathbf{XY}}}{c} \mod 2\pi$$

and c is the speed-of-light.

Then the relative phase difference of these two signals is

$$\begin{aligned}
\phi_{\mathbf{CD}} =& -2\pi \frac{f_{\mathbf{A}} d_{\mathbf{AC}}}{c} + 2\pi \frac{f_{\mathbf{B}} d_{\mathbf{BC}}}{c} + 2\pi \frac{f_{\mathbf{A}} d_{\mathbf{AD}}}{c} - 2\pi \frac{f_{\mathbf{B}} d_{\mathbf{BD}}}{c} \mod 2\pi \\
=& 2\pi \frac{\bar{f}}{c} (d_{\mathbf{AD}} - d_{\mathbf{AC}} + d_{\mathbf{BC}} - d_{\mathbf{BD}}) + \\
& 2\pi \frac{\tilde{f}}{c} (d_{\mathbf{AD}} - d_{\mathbf{AC}} - d_{\mathbf{BC}} + d_{\mathbf{BD}}) \mod 2\pi \\
\simeq& 2\pi \frac{\bar{f}}{c} (d_{\mathbf{AD}} - d_{\mathbf{AC}} + d_{\mathbf{BC}} - d_{\mathbf{BD}}) \mod 2\pi \quad \left(\text{since } \frac{\bar{f}}{c} \simeq 0\right)
\end{aligned} \tag{2.12}$$

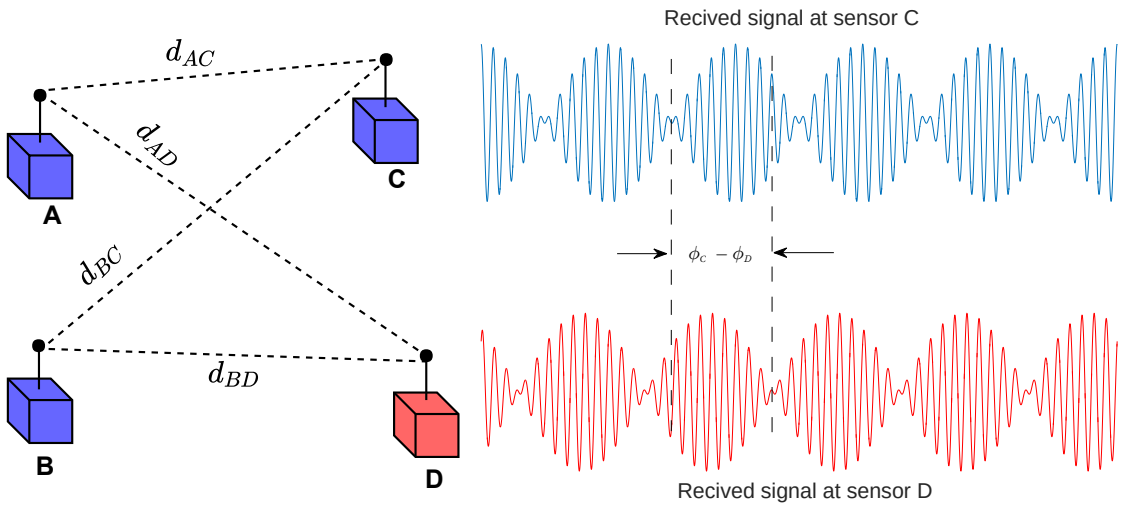


Figure 2.6: The illustrated example of RIPS, where **A**, **B** and **C** are anchors and **D** is the sensor to be localized

We define $\mathbf{ABCD} = d_{\mathbf{AD}} - d_{\mathbf{AC}} + d_{\mathbf{BC}} - d_{\mathbf{BD}}$, then (2.12) becomes

$$\begin{aligned} 2\pi\frac{\bar{f}}{c}d_{\mathbf{ABCD}} \mod 2\pi &= 2\pi\frac{d_{\mathbf{ABCD}}}{\lambda} \mod 2\pi \\ \Longleftrightarrow y_{\mathbf{CD}} &\triangleq \frac{\lambda\phi_{\mathbf{CD}}}{2\pi} = d_{\mathbf{ABCD}} \mod \lambda \qquad (2.13) \\ \Longleftrightarrow d_{\mathbf{ABCD}} &= n_{\mathbf{CD}}\lambda + y_{\mathbf{CD}} \quad \text{where} \quad n_{\mathbf{CD}} \in \mathbb{Z} \qquad (2.14) \end{aligned}$$

where $\bar{f} = \frac{f_{\mathbf{A}}+f_{\mathbf{B}}}{2}$ and $\tilde{f} = \frac{f_{\mathbf{A}}-f_{\mathbf{B}}}{2}$, and $y_{\mathbf{CD}}$ is the equivalent measurement. Then, due to the unknown integer $n_{\mathbf{CD}}$, $d_{\mathbf{ABCD}}$ has a series of candidate values and one of them is the ground truth. This is so-called ambiguity problem.

Similarly, let nodes **A** and **C** be transmitters and nodes **B** and **D** receivers. Then the associated phase difference is

$$\begin{aligned} \phi_{\mathbf{BD}} &= 2\pi\frac{d_{\mathbf{AD}} - d_{\mathbf{CD}} + d_{\mathbf{BC}} - d_{\mathbf{AB}}}{\lambda} \mod 2\pi \\ \Longleftrightarrow y_{\mathbf{BD}} &\triangleq \frac{\lambda\phi_{\mathbf{CD}}}{2\pi} = d_{\mathbf{ABCD}} \mod \lambda \\ \Longleftrightarrow d_{\mathbf{ABCD}} &= n_{\mathbf{BD}}\lambda + y_{\mathbf{BD}} \quad \text{where} \quad n_{\mathbf{BD}} \in \mathbb{Z} \qquad (2.15) \end{aligned}$$

$d_{\mathbf{ABCD}}$ will have different values corresponding to different values of integer $n_{\mathbf{BD}}$. Again

there is an ambiguity problem. A crucial step before implementing the localisation algorithm is to resolve the ambiguous $d_{\mathbf{ABCD}}$ and $d_{\mathbf{ABCD}}$ with measurements $y_{\mathbf{BD}}$ and $y_{\mathbf{CD}}$.

2.2.2 Mathematical Formulation of the Ambiguity Problem

In this section, the rigorous mathematical formulation of ambiguity problem is given and the related properties of ambiguity equations are introduced as well.

Based on the discussion in Section 2.2.1, in general, the ambiguity problem could be written into following two equivalent forms:

$$r = n\lambda + y_{true} \Longleftrightarrow y_{true} = r \mod \lambda \tag{2.16}$$

where $r \in \mathbb{Z}^+$ and $n \in \mathbb{Z}^+$ are unknown values, λ is modular (wavelength used) and y_{true} is remainder (noise-free measurement). The main goal is to solve for r, or equivalently n, via known values y_{true} and λ.

In general, the unknown r is much larger than λ. Then in order to solve the ambiguous r, one should use multiple different wavelengths $\lambda_i > 0$, $i = 1, \cdots, m$, to generate a set of measurements $y_{i,true} > 0$. Therefore, we have a set of ambiguous equations which is called as congruence equations:

$$\begin{cases} r = n_1\lambda_1 + y_{1,true} \\ \vdots \qquad\quad \vdots \\ r = n_m\lambda_m + y_{m,true} \end{cases} \Longleftrightarrow \begin{cases} y_{1,true} = r \mod \lambda_1 \\ \vdots \qquad\quad \vdots \\ y_{m,true} = r \mod \lambda_m \end{cases} \tag{2.17}$$

where $n_i \in \mathbb{Z}$.

Obviously, the equations (2.17) cannot be solved directly as the number of unknown variables is $m + 1$ which is always larger than the number of equations, i.e. m. Therefore, we should resort to other algorithms such as searching algorithm, CRT or lattice algorithm which will be discussed in following sections, to obtain the solution of (2.17).

The unique solution for equations (2.17) is of interest in practice. Therefore, in order to obtain an unique solution for (2.17), there exists a mild condition on the ambiguity resolving algorithms: the unknown distance r is less than the least common multiplier of

the wavelengths used, i.e.

$$r < \mathrm{LCM}(\lambda_1, \cdots, \lambda_m)$$

where $\mathrm{LCM}(\cdot)$ is the least common multiplier function. This condition is justified via Proposition. 2.1 and Lemma 2.2[70]. In general, the $\mathrm{LCM}(\lambda_1, \cdots, \lambda_m)$ is called as the maximum unambiguous range for r which implies that given a set of modulo and remainders, the solution of r is unique within this range.

Proposition 2.1. Suppose that

$$\{z, n_1, \cdots, n_m\} \in \mathbb{Z}^+$$

are unknown and

$$\{\lambda_1, \cdots, \lambda_m\} \in \mathbb{Z}^+, \;\; \lambda_i \neq \lambda_j \;\; \forall i \neq j, \;\; i, j = 1, \cdots, m$$

are known.

Then the solution of following equations

$$\begin{cases} z = & n_1\lambda_1 \\ \vdots & \vdots \\ z = & n_m\lambda_m \end{cases} \tag{2.18}$$

are $z = 0$ or $z = k \cdot \mathrm{LCM}(\lambda_1, \cdots, \lambda_m)$, where $\mathrm{LCM}(\cdot)$ is the least common multiplier function and $k \in \mathbb{Z}$. □

Proof. Generally, let $k = 1$. Suppose that z_0 is a solution of (2.18) and $z_0 \neq 0$, $z_0 \neq \mathrm{LCM}(\lambda_1, \cdots, \lambda_m)$. Therefore, we have

$$n_{1,0}\lambda_1 - n_{2,0}\lambda_2 = z_0 - z_0 = 0$$

From Bézout's identity, the solutions of above equation are

$$\begin{cases} n_{1,0} = 0 \\ n_{2,0} = 0 \end{cases} \quad \text{or} \quad \begin{cases} n_{1,0} = \lambda_2 \\ n_{2,0} = \lambda_1 \end{cases}$$

Similarly, we have following conclusion:

$$\begin{cases} n_{1,0} = 0 \\ n_{2,0} = 0 \\ \vdots \\ n_{m,0} = 0 \end{cases} \quad \text{or} \quad \begin{cases} n_{1,0} = \frac{\text{LCM}(\lambda_1,\cdots,\lambda_m)}{\lambda_1} \\ n_{2,0} = \frac{\text{LCM}(\lambda_1,\cdots,\lambda_m)}{\lambda_2} \\ \vdots \\ n_{m,0} = \frac{\text{LCM}(\lambda_1,\cdots,\lambda_m)}{\lambda_m} \end{cases}$$

Then $z = 0$ or $z = \text{LCM}(\lambda_1, \cdots, \lambda_m)$.

For $k \in \mathbb{Z}$, we have the same result. This completes the proof. □

Then we have following lemma to justify the uniqueness of solution[71].

Lemma 2.2 (The uniqueness of solution)**.** The unknown value (distance) $r \in \mathbb{Z}^+$ can be determined from (2.17) uniquely if and only if $r < \text{LCM}(\lambda_1, \cdots, \lambda_m)$. □

Proof. Suppose that $r_1 < \text{LCM}(\lambda_1, \cdots, \lambda_m)$ and $r_2 < \text{LCM}(\lambda_1, \cdots, \lambda_m)$ solve the equations (2.17) and $r_1 > r_2$. Then we have

$$z = r_1 - r_2 \text{ and } 0 < z < \text{LCM}(\lambda_1, \cdots, \lambda_m)$$

satisfying

$$\begin{cases} z = & (n_{1,1} - n_{2,1})\lambda_1 + y_{1,true} - y_{1,true} = (n_{1,1} - n_{2,1})\lambda_1 \\ \vdots & \vdots \\ z = & (n_{1,m} - n_{2,m})\lambda_m + y_{m,true} - y_{m,true} = (n_{1,m} - n_{2,m})\lambda_m \end{cases} \tag{2.19}$$

where $n_{i,j}$ is the integral solution of j-th equation of r_i, $i = 1, 2$ and $j = 1 \cdots, m$.

Then z only has two possible solution, 0 or $\text{LCM}(\lambda_1, \cdots, \lambda_m)$. However, since $r < \text{LCM}(\lambda_1, \cdots, \lambda_m)$, therefore, $z = 0$ and there exists a unique solution for equations (2.17) if $r < \text{LCM}(\lambda_1, \cdots, \lambda_m)$. □

Accordingly, let $n_{i,max} \in \mathbb{Z}^+$ be the maximum unambiguous value of unknown integer $n_i, i = 1, \cdots, m$ corresponding to the maximum unambiguous range in (2.17), then $n_{i,max}$ could be determined by

$$n_{i,max} = \frac{\text{LCM}(\lambda_1, \cdots, \lambda_m)}{\lambda_i} \tag{2.20}$$

2.3 The Chinese Remainder Theorem(CRT)

2.3.1 The Conventional CRT

The concept of the ambiguity problem and the earliest example of Chinese Remainder Theorem (CRT) are found in ancient Chinese mathematical book *Sun Zi Suanjing* (Sun's Arithmetical Manual), of which the author is unknown. The exact date of this book is also unknown, but it is usually thought to have been written some time between the 3rd and 5th AD. The algorithm was formulated by Aryabhata in 6th century. A more general algorithm, called as *Dayanqiuyi* method was introduced in *Qin Jiushao*'s book *Mathematical Treatise* in 1247. Later, in 1801, Carl Friedrich Gauss discussed congruence equation in his notable book *Disqvisitiones arithmeticae* and invented the notation for congruence equation. The general solution was given by Gauss as well.

The conventional CRT which is used to solve noise-free congruence equations is introduced in this section. Some definitions and notations are introduced before the conventional CRT (Theorem 2.1).

Definition 2.14. Two positive integers λ_1 and λ_2 are said to be co-prime if and only if $\text{GCD}(\lambda_1, \lambda_2) = 1$ and this is denoted by[1][72]

$$\lambda_1 \perp \lambda_2$$

where $\text{GCD}(\cdot)$ is greatest common divisor. □

Definition 2.15. Let $\lambda_1, \lambda_2 \in \mathbb{Z}$, then $b \in \mathbb{Z}$ is said to be the modular inverse of λ_1

[1]This notation is suggested in[72] since "*Like perpendicular lines dont have a common direction, perpendicular numbers dont have common factors*"

modulo λ_2 if

$$b\lambda_1 \text{mod} \quad \lambda_2 = 1$$

□

Definition 2.16. Let $\{b, \lambda_1, \lambda_2\} \in \mathbb{Z}$, λ_1 divides λ_2 if there exists an integer $b \in \mathbb{Z}$ such that $\lambda_1 = b\lambda_2$. If λ_1 divides λ_2, we write $\lambda_1 \mid \lambda_2$, otherwise, $\lambda_1 \nmid \lambda_2$. □

The conventional CRT is encapsulated in the following theorem. The proof can be found in [70].

Theorem 2.1 (Conventional CRT)**.** Consider following congruence system

$$\begin{cases} y_{1,true} = r \mod \lambda_1 \\ \vdots \\ y_{m,true} = r \mod \lambda_m \end{cases} \tag{2.21}$$

where $y_{i,true} \in \mathbb{Z}, i = 1, \cdots, m$ are remainders of $r \in \mathbb{Z}$ modulo $\lambda_i \in \mathbb{Z}$. Assume that $\lambda_i \perp \lambda_j, \forall\ i, j = 1, \cdots, m,\ i \neq j$ and $r < \text{LCM}(\lambda_1, \cdots, \lambda_m)$. Then the unique solution for r exists and is given by

$$r = \sum_{i=1}^{k} \xi_i y_{i,true}$$

where $\xi_i = \frac{\prod_{j=1}^{m} \lambda_j}{\lambda_i} b_i$ and b_i is modular inverse of $\frac{\prod_{j=1}^{m} \lambda_j}{\lambda_i}$ modulo λ_i. □

However, in engineering, the remainders (measurements) has noise and therefore, the conventional CRT cannot be applied directly to estimate the range as it is sensitive to noise[24]. This problem was first investigated by Xia etc. and robust CRT algorithms were proposed in [73, 74] . In [24], a closed-form CRT is shown to be robust to noise and computationally efficient. Later, some improved algorithms are presented in [75–78]. In order to increase the probability of successful reconstruction of the ambiguous number in the context of noise, an improved algorithm based on the modified CRT is presented in next section. It is demonstrated that the improved algorithm outperforms the traditional closed-form CRT in terms of probability of successful reconstruction.

Algorithm 2.1: LLL basis reduction algorithm[57]

Data: A lattice basis $[\mathbf{a}_1, \cdots, \mathbf{a}_m]$ and $\frac{1}{4} < \delta < 1$
Result: Reduced basis $[\mathbf{b}_1, \cdots, \mathbf{b}_m]$

Main Loop
- **for** $i = 1 : m$ **do** $\mathbf{b}_i = \mathbf{a}_i$
- **for** $i = 1 : m$ **do**
  - $\mathbf{a}_i^* = \mathbf{a}_i$;
  - **for** $j = 1 : i - 1$ **do** $\mu_{i,j} = (\mathbf{b}_i \cdot \mathbf{b}_j^*)/\gamma_i$ and $\mathbf{b}_i^* = \mathbf{b}_i^* - \mu_{i,j}\mathbf{b}_j^*$;
  - $\gamma_i = \mathbf{b}_i^* \cdot \mathbf{b}_i^*$;
- **end**
- $k = 2$;
- **while** $k \leq n$ **do**
  - Call `reduce` $(k, k-1)$;
  - **if** $\gamma_k \geq (\delta - \mu_{k,k-1}^2)\gamma_{k-1}$ **then**
    - **for** $j = k - 2 : -1 : 1$ **do** Call `reduce` (k, j) ;
    - $k = k + 1$;
  - **else**
    - Call `swap` (k) ;
    - **if** $k > 2$ **then** $k = k - 1$
  - **end**
- **end**

Procedure `reduce` (k, j)
- **if** $|\mu_{k,j}| > \frac{1}{2}$ **then**
  - $\mathbf{a}_k = \mathbf{A}_k - \lceil \mu_{k,j} \rfloor \mathbf{a}_j$;
  - **for** $i = 1 : j - 1$ **do** $\mu_{k,i} = \mu_{k,i} - \lceil \mu_{k,j} \rfloor \mu_{j,i}$;
  - $\mu_{k,j} = \mu_{k,j} - \lceil \mu_{k,j} \rfloor$
- **end**

Procedure `swap` (k)
- $z = \mathbf{b}_{k-1}$, $\mathbf{b}_{k-1} = \mathbf{b}_k$, $\mathbf{b}_k = z$;
- $\nu = \mu_{k,k-1}$, $\eta = \gamma_k + \eta^2\gamma_{k-1}$;
- $\mu_{k,k-1} = \nu\gamma_{k-1}/\eta$, $\gamma_k = \gamma_k\gamma_{k-1}/\eta$, $\gamma_{k-1} = \eta$;
- **for** $j = 1 : k - 2$ **do** $t = \mu_{k-1,j}$, $\mu_{k-1,j} = \mu_{k,j}$, $\mu_{k,j} = t$;
- **for** $i = (k + 1) : m$ **do**
  - $\xi = \mu_{i,k}$, $\mu_{i,k} = \mu_{i,k-1} - \nu\mu_{i,k}$, $\mu_{i,k-1} = \mu_{k,k-1}\mu_{i,k} + \xi$
- **end**

Algorithm 2.3: Sphere Decoder

Data: The $m \times m$ upper triangular matrix $\mathbf{R}$ with positive diagonal elements and an m-dimensional vector $\boldsymbol{x} \in \mathbb{R}^m$
Result: Integer vector $\boldsymbol{n}$

$\boldsymbol{H} = \mathbf{R}^{-1}$
$\boldsymbol{e} = \mathbf{0}_{m \times m}$
$bestdist = \infty$
$k = 1$
$d_k = 0$
$\boldsymbol{e}_k = \boldsymbol{x}\boldsymbol{H}$
$u_k = \lfloor e_{k,k} \rceil$
$y = \frac{e_{k,k} - u_k}{h_{k,k}}$
$s_k = sgn(y)$
while $Stopconditionisnotsatisfied$ **do**
  $newdist = d_k + y^2$
  if $newdist < bestdist$ **then**
    if $k \neq m$ **then**
      for $i = K + 1 : m$ **do** $e_{k+1,i} = e_{k,i} - y h_{k,i}$;
      $k = k + 1$
      $d_k = newdist$
      $u_k = \lfloor e_{k,k} \rceil$
      $y = \frac{e_{k,k} - u_k}{h_{k,k}}$
      $s_k = sgn(y)$
    else
      $\boldsymbol{n} = \mathbf{U}$
      $bestdist = newdist$
      $k = k - 1$
      $u_k = u_k + s_k$
      $y = \frac{e_{k,k} - u_k}{h_{k,k}}$
      $s_k = -s_k - sgn(s_k)$
  else
    if $k = 1$ **then**
      return $\boldsymbol{n}$
    else
      $k = k - 1$
      $u_k = u_k + s_k$
      $y = \frac{e_{k,k} - u_k}{h_{k,k}}$
      $s_k = -s_k - sgn(s_k)$

Chapter 3

Efficient Algorithms to Resolve Ambiguity in Noise — CRT

The well-known CRT is a powerful tool to solve the congruence equations arising in mathematics and engineering. In [24], a robust closed-form CRT is presented to resolve the congruence equations with noisy remainders(measurements). Some of the improvements of the closed-form CRT are proposed in order to extend the applications of this algorithm[75–78]. In this chapter, we aim to introduce the modified CRT[76] and prove the related lemmas, then we present an improved CRT algorithm based on the modified CRT using the noise property which has better performance in terms of reconstruction probability. The structure of this chapter is as follows: the noisy ambiguous measurement model is firstly introduced in Section 3.1; in Section 3.2 the modified CRT algorithm with noisy measurements is described and an improved CRT method is presented in order to increase the reconstruction probability in Section 3.2. Finally, Section 3.4 concludes this chapter.

3.1 The Measurement Model

As introduced in the last chapter, in sensor localization, the measurement is normally in the form of phase and the distance of interest can be calculated from

$$\phi_{true} = 2\pi\frac{r}{\lambda} \mod 2\pi \tag{3.1}$$

where ϕ_{true} is noise-free phase measurement, λ is used wavelength, r is an unknown distance.

In practice, the measurement is always contaminated by noise, i.e.

$$\phi = \phi_{true} + \omega^{\phi} \tag{3.2}$$

where ω^{ϕ} is the phase noise term, assumed to be Normally distributed with mean 0 and variance δ'^2, i.e. $\omega^{\phi} \sim \mathcal{N}(0, \delta'^2)$. Therefore, the noisy form of (3.1) is

$$\phi_{true} + \omega^{\phi} = 2\pi\frac{r}{\lambda} \mod 2\pi + \omega^{\phi} \tag{3.3}$$

which can be rewritten as

$$2\pi\lambda\phi_{true} + 2\pi\lambda\omega^{\phi} = r \mod \lambda + 2\pi\lambda\omega^{\phi} \tag{3.4}$$

$$\Longleftrightarrow y_{true} + \omega = r \mod \lambda + \omega \tag{3.5}$$

where $\omega = 2\pi\lambda\omega^{\phi}$ has Gaussian distribution with mean 0 and variance $4\pi\lambda^2\delta'^2$, denoted $4\pi\delta'^2$ by δ^2, since $\omega^{\phi} \sim \mathcal{N}(0, \delta'^2)$. The we have $\omega \sim \mathcal{N}(0, \delta^2\lambda^2)$.

Equivalently, (3.5) can be written into:

$$c = n\lambda + y_{true} + \omega \Longrightarrow c = n\lambda + y \tag{3.6}$$

where c is the possible distance due to the noise, $n \in \mathbb{Z}$ and $y = y_{true} + \omega$.

In applications, the available wavelengths are normally much smaller than the required unambiguous range, so that multiple measurements, $\{y_1, \cdots, y_m\}$, with different wavelengths, $\{\lambda_1, \cdots, \lambda_m\}$, are required to provide an unambiguous estimate of the distance r. We then have following congruence system:

$$\begin{cases} c_1 = & n_1\lambda_1 + y_{1,true} + \omega_1 \\ \vdots & \\ c_m = & n_m\lambda_m + y_{m,true} + \omega_m \end{cases} \Longrightarrow \begin{cases} c_1 = & n_1\lambda_1 + y_1 \\ \vdots & \\ c_m = & n_m\lambda_m + y_m \end{cases} \tag{3.7}$$

where $\omega_i \sim \mathcal{N}(0, \lambda_i^2\delta^2)$.

3.2 The Modified CRT

The conventional CRT described in Section 2.3 is inapplicable since is not suitable to find a robust estimate in a noisy context. In this section, a modified CRT is introduced which

could be used in noisy case. The modified CRT is described in Theorem 3.1 [24,76] and some propositions and lemma are proved as well.

Proposition 3.1. Let $q \in \mathbb{Z}^-$ and $\lambda \in \mathbb{Z}^+$, then $q \mod \lambda$ can be calculated by

$$q \mod \lambda = \lambda - (|q| \mod \lambda).$$

□

Proof. Let $y^* = |q| \mod \lambda$, then there exist an $N \in \mathbb{Z}^+$, such that $|q| = N\lambda + y^*$. Therefore,

$$\begin{aligned} y^* =& |q| - N\lambda \\ \lambda - y^* =& \lambda - (|q| - N\lambda) \\ -|q| =& -(1+N)\lambda + (\lambda - y^*) \end{aligned}$$

Then $0 < (\lambda - y^*) < \lambda$ is the remainder of the q modulo λ. □

Proposition 3.2. Let $\{\lambda_1, \cdots, \lambda_m\} \in \mathbb{Z}^+$, and $\bar{\lambda}_i \perp \bar{\lambda}_j, \ \forall\ i, j = 1, \cdots, m-1,\ i \neq j$, where $\bar{\lambda}_i = \frac{\lambda_i}{g_{i,m}}$ and $g_{i,m} \triangleq \text{GCD}(\lambda_i, \lambda_m)$, where $\text{GCD}(\cdot)$ is greatest common divisor. Then

$$\text{LCM}(\lambda_1, \cdots, \lambda_m) = \lambda_m \prod_{i=1}^{m-1} \bar{\lambda}_i$$

□

Proof. Let $m = 2$, then it is easily to see that $\text{LCM}(\lambda_1, \lambda_2) = \frac{\lambda_1 \lambda_2}{g_{1,2}} = \bar{\lambda}_1 \lambda_2$. Let $m = 3$, then we have

$$\begin{aligned} \text{LCM}(\lambda_1, \lambda_2, \lambda_3) &= \text{LCM}(\text{LCM}(\lambda_1, \lambda_3), \lambda_2) \\ &= \text{LCM}(\bar{\lambda}_1 \lambda_3, \lambda_2) \\ &= g_{2,3} \text{LCM}\left(\bar{\lambda}_1 \frac{\lambda_3}{g_{2,3}}, \frac{\lambda_2}{g_{2,3}}\right) \\ &= g_{2,3} \bar{\lambda}_1 \frac{\lambda_3}{g_{2,3}} \frac{\lambda_2}{g_{2,3}} = \bar{\lambda}_1 \bar{\lambda}_2 \lambda_3 \end{aligned}$$

Repeating the same steps, we have $\text{LCM}(\lambda_1, \cdots, \lambda_m) = \lambda_m \prod_{i=1}^{m-1} \bar{\lambda}_i$ □

Lemma 3.1 is crucial in proving Theorem 3.1. A slightly different lemma is given in [24], we will indicate the difference between these two lemmas in the subsequent remark.

Lemma 3.1. Let $\{\lambda_1, \lambda_2, x, y\} \in \mathbb{Z}^+$ and $q \in \mathbb{Z}$, where λ_1, λ_2, q are known and x, y are unknown, furthermore, suppose $\lambda_1 \perp \lambda_2$. Consider following equations,

$$\lambda_1 x - \lambda_2 y = q \tag{3.8}$$

then the integral solution for $\{x, y\}$ is, $n \in \mathbb{Z}$,

$$\begin{cases} x = (bq) \mod \lambda_2 + \lambda_2 n \\ y = \frac{((bq) \mod \lambda_2 + \lambda_2 n)\lambda_1 - q}{\lambda_2} \end{cases} \tag{3.9}$$

where $b \in \mathbb{Z}^+$ is the modular inverse of λ_1 modulo λ_2. □

Proof. From Bézout's lemma[79], we know that there always exists a modular inverse of λ_1 modulo λ_2 if $\lambda_1 \perp \lambda_2$, i.e.

$$\lambda_1 b - \lambda_2 \frac{\lambda_1 b - 1}{\lambda_2} = 1$$
$$\lambda_1 bq - \lambda_2 \frac{\lambda_1 b - 1}{\lambda_2} q = q$$

where $\frac{\lambda_1 b - 1}{\lambda_2} \in \mathbb{Z}$.

If $bq \geq 0$, then $x = bq$ is a specific solution of (3.8) and

$$x = (bq) \mod \lambda_2$$

is also a solution of (3.8) since $\exists N \in \mathbb{Z}$ such that

$$\begin{aligned} y &= \frac{\lambda_1((bq) \mod \lambda_2) - q}{\lambda_2} \\ &= \frac{\lambda_1(bq - N\lambda_2) - q}{\lambda_2} \in \mathbb{Z} \end{aligned}$$

If $bq < 0$, we have a similar result using Prop.3.1. This implies that $(bq) \mod \lambda_2$ is a specific solution of (3.8) and the general solution is (3.9).

Next, we aim to show that all the solutions of (3.8) are included in (3.9). Suppose that there exits an integral solution $\{x_0, y_0\}$ for (3.8) which does not satisfy (3.9). Then, for specific $\{x_0, y_0\}$, one can find $\alpha \in \mathbb{Z}$ and $n_0 \in \mathbb{Z}$ such that

$$\begin{cases} x_0 = (bq) \mod \lambda_2 + \lambda_2 n_0 + \alpha \\ y_0 = \frac{((bq) \mod \lambda_2 + \lambda_2 n_0 + \alpha)\lambda_1 - q}{\lambda_2} \end{cases} \tag{3.10}$$

where $\alpha \neq 0$ and $\alpha \nmid \lambda_2$. Obviously,

$$\begin{aligned} y_0 =& \frac{((bq) \mod \lambda_2 + \lambda_2 n_0 + \alpha)\lambda_1 - q}{\lambda_2} \\ =& \frac{((bq) \mod \lambda_2 + \lambda_2 n_0)\lambda_1 - q}{\lambda_2} + \frac{\alpha\lambda_1}{\lambda_2} \\ \triangleq& C + \frac{\alpha\lambda_1}{\lambda_2} \end{aligned}$$

where $C \in \mathbb{Z}$ based on previous discussion. Since $y_0 \in \mathbb{Z}$, therefore α is either 0 or $\alpha \mid \lambda_2$ which contradicts the assumption. Therefore, all solutions are included in (3.9). □

Remark 3.1. Another solution with slightly different form to (3.8) is given in [24](see Lemma 1, [24]) as:

$$\begin{cases} x = bq + \lambda_2 nq \\ y = \frac{(bq + \lambda_2 nq)\lambda_1 - q}{\lambda_2} \end{cases} \tag{3.11}$$

where b is the modular inverse of λ_1 modulo λ_2. However, (3.11) dose not give all solutions of (3.8). For example, consider an equation

$$7x - 9y = 2$$

with specific solution $\{x = 17, y = 13\}$. Then we can calculate that $b = 4$ is the modular

inverse of 7 modulo 9. Therefore, from (3.11), the solution of $7x - 9y = 2$ is given by

$$x = 4 \cdot 2 + 9 \cdot 2 \cdot n \neq 17, \quad \forall n \in \mathbb{Z}.$$

This because of that the term nq in (3.11) belongs to

$$\{z | z \in \mathbb{Z} \text{ and } z \mod q \equiv 0\}$$

rather than $\mathbb{Z}$. □

Recall the congruence system

$$\begin{cases} y_{1,true} = r \mod \lambda_1 \\ \vdots \\ y_{m,true} = r \mod \lambda_m \end{cases} \tag{3.12}$$

Based on conventional CRT and Lemma 3.1, the modified CRT [76] is given in Theorem 3.1.

Theorem 3.1. Consider the congruence system shown in (3.12). Assume that

$$r < \text{LCM}(\lambda_1, \cdots, \lambda_m) \text{ and } \bar{\lambda}_i \perp \bar{\lambda}_j, \ \forall\, i, j = 1, \cdots, m-1, \ i \neq j$$

where $\bar{\lambda}_i = \frac{\lambda_i}{g_{i,m}}$ and $g_{i,m} \triangleq \text{GCD}(\lambda_i, \lambda_m)$, $\forall i = 1, \cdots, m-1$. Then there is a unique solution for r. □

Proof. The system (3.12) can be rewritten into

$$\begin{cases} n_1\lambda_1 + y_{1,true} = r \\ \vdots \\ n_m\lambda_m + y_{m,true} = r \end{cases} \tag{3.13}$$

where $n_1, \cdots, n_m$ are unknown integers.

Then we subtract the last equation in (3.13) from the first $m-1$ equations, i.e.

$$\begin{cases} n_1\lambda_1 + y_{1,true} - (n_m\lambda_m + y_{m,true}) = 0 \\ \vdots \\ n_{m-1}\lambda_{m-1} + y_{m-1,true} - (n_m\lambda_m + y_{m,true}) = 0 \end{cases}$$
$$\Longrightarrow \begin{cases} n_m\lambda_m - n_1\lambda_1 & = y_{m,true} - y_1 & \triangleq q_{1,true} \\ \vdots & & \\ n_m\lambda_m - n_{m-1}\lambda_{m-1} & = y_{m,true} - y_{m-1,true} & \triangleq q_{m-1,true} \end{cases}$$

Consider $g_{i,m} = \text{GCD}(\lambda_i, \lambda_m)$ for $i = 1, \cdots, m-1$, we have

$$\begin{cases} n_m\frac{\lambda_m}{g_{1,m}} - n_1\frac{\lambda_1}{g_{1,m}} & = \frac{q_{1,true}}{g_{1,m}} \\ \vdots & \\ n_m\frac{\lambda_m}{g_{m-1,m}} - n_{m-1}\frac{\lambda_m - 1}{g_{m-1,m}} & = \frac{q_{m-1,true}}{g_{m-1,m}} \end{cases}$$
$$\Longrightarrow \begin{cases} n_m\bar{\lambda}_m - n_1\bar{\lambda}_1 & = \bar{q}_{1,true} \\ \vdots & \\ n_m\bar{\lambda}_m - n_{m-1}\bar{\lambda}_{m-1} & = \bar{q}_{m-1,true} \end{cases} \tag{3.14}$$

where $\bar{\lambda}_{i,m} = \frac{\lambda_m}{g_{i,m}}$ and $\bar{q}_{i,true} = \frac{q_{i,true}}{g_{i,m}}$ for $i = 1, \cdots, m-1$. Since all numbers in $n_m\bar{\lambda}_m - n_1\bar{\lambda}_1$ are integers, $\bar{q}_{i,true} \in \mathbb{Z}$.

From Lemma 3.1 and assumption $\bar{\lambda}_i \perp \bar{\lambda}_j,\ \forall\, i, j = 1, \cdots, m,\ i \neq j$ we know that the solutions for the unknown integer n_m in (3.14) is obtainable as follows:

$$\begin{cases} n_m = & (b_1\bar{q}_{1,true}) \mod \bar{\lambda}_1 + n_1\bar{\lambda}_1 \\ \vdots & \\ n_m = & (b_{m-1}\bar{q}_{m-1,true}) \mod \bar{\lambda}_{m-1} + n_{m-1}\bar{\lambda}_{m-1} \end{cases}$$
$$\Longrightarrow \begin{cases} n_m = & \bar{y}_{1,true} + n_1\bar{\lambda}_1 \\ \vdots & \\ n_m = & \bar{y}_{m-1,true} + n_{m-1}\bar{\lambda}_{m-1} \end{cases} \tag{3.15}$$

where $\bar{y}_i = (b_i\bar{q}_i) \mod \bar{\lambda}_i$ and $0 \leq \bar{y}_i < \bar{\lambda}_i, i = 1, \cdots, m-1$. System (3.15) is a con-

gruence system with co-prime moduli $\{\lambda_1, \cdots, \lambda_{m-1}\}$ and remainders $\{y_1, \cdots, y_{m-1}\}$. From Theorem 3.15, we know that the solution of n_m is

$$n_m = \sum_{i=1}^{m-1} \bar{\xi}_i \bar{y}_i$$

where $\bar{\xi}_i = \frac{\prod_{j=1}^{m} \bar{\lambda}_j}{\bar{\lambda}_i} \bar{b}_i$ and $\bar{\xi}_i \mod \bar{\lambda}_i = 1$. Then we have $r = n_m \lambda_m + y_m$.

From (3.15) and Prop. 3.2, we know that $n_m < \prod_{i=1}^{m-1} \bar{\lambda}_i$, meanwhile, from CRT, $n_m < \text{LCM}(\bar{\lambda}_1, \cdots, \bar{\lambda}_{m-1}) = \prod_{i=1}^{m-1} \bar{\lambda}_i$. Therefore, n_m is the unique solution for n_m for system (3.13) and r has a unique solution. □

In the next section, we will show how to use Theorem 3.1 to solve the noisy distance problem and present an improved CRT algorithm.

3.3 An Improved CRT Solution

3.3.1 Introduction

In the noise-free scenario, the conventional CRT is enough to find the ambiguous integer effectively provided the associated conditions are satisfied. However, in the noisy case, this theorem cannot be applied directly as it is sensitive to noise, and will give the incorrect answer with high probability. Therefore it is important to find robust algorithms that are less sensitive to the noise. It is shown that Theorem 3.1 can be modified to solve noisy ambiguity problem efficiently.

In practice, estimating $\{\bar{q}_{1,true}, \cdots, \bar{q}_{m-1,true}\}$ in (3.16) and (3.14) is a key step in Theorem 3.1 for estimating r via the noisy measurements $\{y_1, \cdots, y_m\}$ which are not integers. Define, for $i = 1, \cdots, m-1$,

$$\bar{q}_i \triangleq \frac{(y_i - y_m)}{g_{i,m}}, \text{ and } \bar{q}_{i,true} \triangleq \frac{(y_{i,true} - y_{m,true})}{g_{i,m}}. \tag{3.16}$$

Since $\bar{q}_i$ are no longer integers so that the modified CRT is inapplicable. As suggested in [76], a rounding operation $\hat{\bar{q}}_i = \lceil \bar{q}_i \rfloor$ is introduced for the estimation of $\bar{q}_{i,ture}$. Theorem

3.1 is then applicable to obtain the estimation of n_m, denoted by $\hat{n}_m$. An improved CRT algorithm is presented in next section.

3.3.2 The Proposed CRT Algorithm

Let the set of waveforms used be $\{\lambda_1, \cdots, \lambda_m\}$, and suppose that they satisfy the conditions of Theorem 3.1. Also assume that $\lambda_1 > \lambda_2 > \cdots > \lambda_m$. The smallest wavelength λ_m is selected as the "reference" in the measurement operation (3.14) because ω_m has the minimum variance as $\omega_i \sim \mathcal{N}(0, \delta^2\lambda_i)$.

After obtaining $\hat{n}_m$, the estimation of r, $\hat{r}$, can be done using a maximum likelihood method [27], i.e

$$\hat{r} = W \sum_{i=1}^{m} (\hat{n}_i \lambda_i + y_i) W_i \tag{3.17}$$

where $W_i = 1/\lambda_i^2$, $W = 1/\sum_{i=1}^{m} W_i$ and $\hat{n}_i = \left\lceil \frac{\hat{n}_m \lambda_m + y_m - y_i}{\lambda_i} \right\rfloor$.

In this estimation, the calculation of $\hat{\tilde{q}}_i$ is crucial and the algorithm will return the correct answer if[76]

$$-1/2 < \frac{\omega_i - \omega_m}{g_{i,m}} < 1/2, \quad \forall i = 1, \cdots, m-1$$

otherwise, the algorithm will return the wrong $\hat{n}_m$. Furthermore,

$$\left[\frac{\omega_1 - \omega_m}{g_{1,m}}, \cdots, \frac{\omega_{m-1} - \omega_m}{g_{m-1,m}} \right] \sim \mathcal{N}\left(\mathbf{0}, \delta^2 \mathbf{\Sigma}\right)$$

where $\mathbf{0}$ is zero vector and $\delta^2\mathbf{\Sigma}$ is covariance matrix. Since $\left[\frac{\omega_1 - \omega_m}{g_{1,m}}, \cdots, \frac{\omega_{m-1} - \omega_m}{g_{m-1,m}} \right]$ can be written into $[\omega_1, \cdots, \omega_m]\,\mathbf{M}^T$ and $[\omega_1, \cdots, \omega_m]$ is Normally distributed with mean 0 and covariance

$$\delta^2 \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ 0 & 0 & \cdots & \lambda_m^2 \end{bmatrix}$$

where

$$\mathbf{M} = \begin{bmatrix} \frac{1}{g_{1,m}} & 0 & \cdots & 0 & -\frac{1}{g_{1,m}} \\ 0 & \frac{1}{g_{2,m}} & \cdots & 0 & -\frac{1}{g_{2,m}} \\ 0 & 0 & \cdots & \frac{1}{g_{m-1,m}} & \frac{1}{g_{m-1,m}} \end{bmatrix}$$

then we have

$$\mathbf{\Sigma} = \delta^2 \mathbf{M} \begin{bmatrix} \lambda_1^2 & 0 & \cdots & 0 \\ 0 & \lambda_2^2 & \cdots & 0 \\ 0 & 0 & \cdots & \lambda_m^2 \end{bmatrix} \mathbf{M}^T$$

This CRT solution using rounding operation implies that there exists an integer set $\mathbf{Z}_0 = [z_1, \cdots, z_{m-1}]$ such that $\hat{\bar{q}}_i + z_i = \bar{q}_{i,true},\ \forall i = 1, \cdots, m-1$ for a given set of measurement $\{y_1, \cdots, y_m\}$. Let $\hat{\bar{\mathbf{q}}} = [\hat{\bar{q}}_1, \cdots, \hat{\bar{q}}_{m-1}]$. Then the algorithm implemented under Theorem 3.1 will return the correct $\hat{n}_m$, i.e. $\hat{n}_m = n_m$, if $\hat{\bar{\mathbf{q}}} + \mathbf{Z}_0$ is used. However, in practice, the elements contained in $\mathbf{Z}_0$ are unknown, therefore we need to estimate the integer set $\mathbf{Z}_0$.

Let $\mathbf{Z}_j = [z_{j,1}, \cdots, z_{j,m-1}],\ j = 1, 2, \cdots,\ z_{j,i} \in \mathbb{Z}$ represent possible values of $\mathbf{Z}_0$. In practice, since the δ is small and $\lambda_m < \lambda_i, \forall i = 1, \cdots, m-1$, the variance of $(\omega_i - \omega_m)$ is not large. We may reasonably assume that $z_{j,i} \in \{-1, 0, 1\}$[1]. It is easy to find the max value of j is $j_{max} = 3^{m-1}$. Let $\mathcal{Z} = [\mathbf{Z}_1^T, \cdots, \mathbf{Z}_{j_{max}}^T]^T$ be a $j_{max} \times (m-1)$ dimensional matrix containing all possible value of $\mathbf{Z}_0$. The probability that $\mathbf{Z}_0 \in \mathcal{Z}$ is calculated by

$$Pr\left(\mathbf{Z}_0 \in \mathcal{Z}\right) = \int_{\mathbf{\Omega}} \frac{1}{\left(2\pi\delta^2 |\mathbf{\Sigma}|\right)^{\frac{m}{2}}} \exp\left\{-\frac{1}{2}\mathbf{x}^T (\delta^2 \mathbf{\Sigma})^{-1} \mathbf{x}\right\} d\mathbf{x}$$

where $\mathbf{\Omega}$ is the integration volume which is a (hyper-) rectangle with lower and upper limits $-\frac{3}{2}$ and $\frac{3}{2}$ respectively.

Now the main goal is to find the estimation of $\mathbf{Z}_0$. An efficient algorithm is presented

[1] In the standard CRT algorithm, it is assume that $z_{j,i} = 0, \forall i = 1, \cdots, m-1$

as follows:

$$\hat{j}_0 = \arg \min_{j \in \{1,\cdots,3^{m-1}\}} \frac{1}{m} \sum_{i=1}^{m} \hat{\omega}_{j,i}^2 \tag{3.18}$$

$$s.t. \quad \hat{n}_{j,m} = \mathbf{mCRT}\left(\hat{\mathbf{q}} + \mathbf{Z}_j\right), \ \ \mathbf{Z}_j \in \mathcal{Z} \tag{3.19}$$

$$\hat{n}_{j,i} = \left\lceil \frac{\hat{n}_{j,m}\lambda_m + y_m - y_i}{\lambda_i} \right\rceil, \forall i = 1, \cdots, m-1 \tag{3.20}$$

$$\hat{\omega}_{j,i} = \frac{\lfloor \hat{r}_j \rceil - (\hat{n}_{j,i}\lambda_i + y_i)}{\lambda_i} \tag{3.21}$$

$$\hat{r}_j = \mathbf{E}[r|\hat{n}_{j,1}, \cdots, \hat{n}_{j,m}] \tag{3.22}$$

where $\mathbf{E}[r|\hat{n}_{j,1}, \cdots, \hat{n}_{j,m}]$ is the estimation of r given $\hat{n}_{j,1}, \cdots, \hat{n}_{j,m}$ using (3.17), **mCRT** is the modified CRT algorithm described in Theorem 3.1. Thus, the estimation of $\mathbf{Z}_0$ is the $\hat{j}_0$-th row of $\mathcal{Z}$ and, therefore, the new estimation of n_m is $\hat{n}_{\hat{j}_0,m}$, i.e. $\hat{n}_{\hat{j}_0,m} = \mathbf{mCRT}(\hat{\mathbf{q}} + \mathbf{Z}_{\hat{j}_0})$.

We briefly investigate the performance of estimator (3.18) and the conclusion (3.23) is given without proof. The detailed analysis is given in Section 4.5.3. In the proposed algorithm (3.18), formula (3.19) and (3.20) estimate the $\left[\hat{n}_{j,1}, \cdots, \hat{n}_{j,m}\right]$ according to vector $\mathbf{Z}_j$ and fixed $\hat{\mathbf{q}}$, and (3.21) estimates the measurements noise base on $\left[\hat{n}_{j,1}, \cdots, \hat{n}_{j,m}\right]$. If $\mathbf{Z}_0$ is in $\mathcal{Z}$, then the $\hat{r}_j$ is close to true r and the corresponding noise estimation $\left[\hat{\omega}_{j,1}, \cdots, \hat{\omega}_{j,m}\right]$ is approximately Normally distributed with mean 0 and δ^2. Furthermore, if the frequencies used are from same band, then they will close to each other and we have following conclusion: the reconstruction probability given δ^2, $Pr(\hat{n}_{\hat{j}_0,m} = n_m|\delta^2)$, of the proposed algorithm (3.18) can be approximated by

$$\prod_{i=1}^{k} Pr\left(-\frac{1}{\tilde{\lambda}} + 2|\tilde{\omega}_i| < 0\right) \tag{3.23}$$

where $\tilde{\lambda} = \frac{1}{m}\sum_{i=1}^{m} \lambda_i$ and $\tilde{\omega} \sim \mathcal{N}(0, \delta^2)$. (3.23) will be discussed in detail in Section 4.5.

Meanwhile, it can be noticed, in (3.18), the proposed algorithm needs to execute the conventional one 3^{m-1} times to find the required estimation, however, in practice, the number of wavelengths used is not too large and the conventional algorithm is extremely fast, therefore the proposed algorithm will not increase too much computational cost.

3.3.3 Simulation

The proposed algorithm in Section 3.3.2 is evaluated via simulation using two sets of wavelengths:

$$\mathbf{\Lambda}_1 = \{61,\ 59,\ 49,\ 47,\ 46\}$$

and

$$\mathbf{\Lambda}_2 = \{21,\ 19,\ 17,\ 13\}$$

The first set of wavelengths is from Wifi 5 GHz band and the latter from $1 \sim 26$MHz which is from frequency difference region of US UHF band, $902 \sim 928$MHz[32]. These two bands are widely used in localization via range estimation. The distance r is randomly selected between 0 and the LCM of all wavelengths used. The parameter δ in the measurement noise variance $\delta^2\lambda_i^2$ is chosen as $-20\log_{10}\delta = 46 : 2 : 66$, which provides an indication for noise level in the simulation. All simulation results illustrated are averaged over 5000 Monte Carlo runs. The algorithm performance is evaluated in terms of the reconstruction probability $Pr(n_m = \hat{n}_m)$ versus the phase measurement noise level.

The simulation results of the reconstruction probability via Monte Carlo simulation and theoretic computation via (3.23) are given in Fig.3.1. It is found that the simulation results agree with the theoretic result. For example, in the $\mathbf{\Lambda}_1$ case, as shown in the figure, we can notice that the simulated reconstruction probability of ambiguous integers, denoted by blue circle line, agrees with the computed reconstruction probability using (3.23), denoted by star markers, at each error levels. As comparison, the performance of existing algorithm **mCRT** is plotted along with that of the proposed algorithm in Fig.3.1. Clearly, using both wavelength sets the proposed algorithm outperforms **mCRT** in terms of reconstruction probability of the underlying integer set when signal to noise ratio is low.

Finally, we point out that additional computational overhead of the proposed algorithm over the existing CRT algorithm in [24] is marginal.

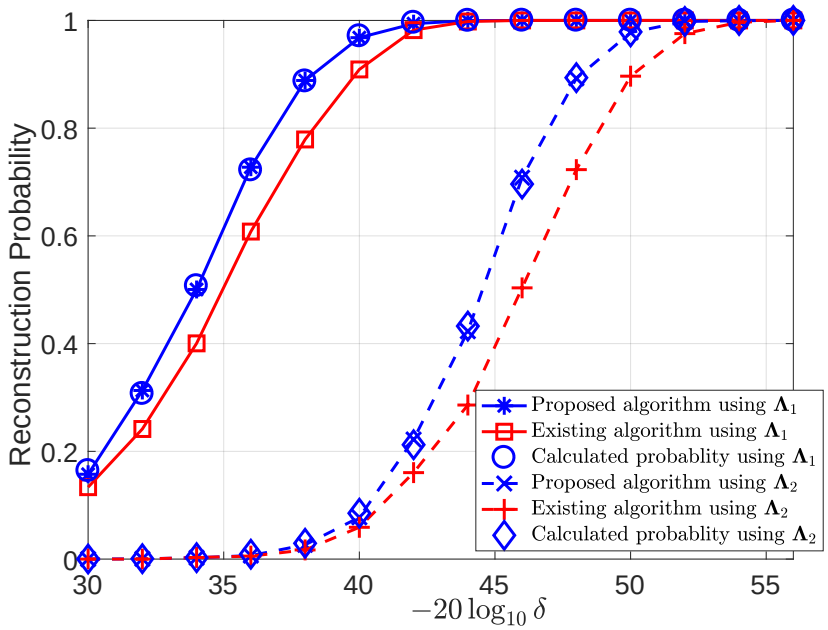


Figure 3.1: Reconstruction probability versus error level

3.4 Conclusions

In this chapter, the modified CRT for noisy ambiguity problem is described and an improved CRT algorithm for range estimation is proposed. Compared with existing approaches, the proposed algorithm takes a set of more probable integer outcomes after the rounding operation is taken into evaluating additional hypotheses based on measurements and known conditions. The approximate probability of correct integer set reconstruction is accordingly given. Simulation results are presented showing the performance of the proposed algorithm is better than the conventional one in terms of reconstruction probability with a slightly higher computations.

Chapter 4

Efficient Algorithms to Resolve Ambiguity in Noise — Lattice Algorithms

In [27], it is shown that ambiguity problem can be formulated in the framework of lattice theory and an efficient algorithm is presented with the same performance as closed-form CRT algorithm. In this chapter, some improved algorithms based on conventional lattice method are proposed. The conventional algorithm is described in Section 4.1; in Section 4.2, a relaxed condition on the selection of wavelengths is presented that extends the applications of the lattice; in Section 4.3, the ambiguity problem with non-coprime wavelengths is addressed using an improved iterative lattice algorithm. A new lattice algorithm with higher reconstruction probability using different basis structures is proposed in Section 4.4. In Section 4.5, an improvement to the algorithm that takes into account multiple lattice points is proposed. This algorithm has better reconstruction probability over the conventional algorithm for a given basis.

4.1 Formulation of the Ambiguity Problem in Lattices

4.1.1 Introduction

In this section, the lattice based algorithm presented in [27] is introduced. The main focus is to elucidate how to transform the phase ambiguity problem into a nearest point problem in the lattice theory. Maximum likelihood estimation of the ambiguous distance from phase measurements is derived. Then a closed-form algorithm based on the lattice method and number theory is described. The performance of the lattice algorithm is analysed.

Recall the ambiguous equations (3.6):

$$\begin{cases} c_1 = & n_1\lambda_1 + y_1 \\ \vdots & \\ c_m = & n_m\lambda_m + y_m \end{cases} \tag{4.1}$$

Since $y_i \sim \mathcal{N}(y_{i,true}, \lambda_i^2\delta^2)$, then $c_i \sim \mathcal{N}(n_i\lambda_i + y_{i,true}, \lambda_i^2\delta^2)$. Denote the true solution of n_i by N_i and let $\mathbf{N} = [N_1, \cdots, N_m]$.

The associated likelihood function is defined by

$$\begin{aligned} f(y_i|r) &= \frac{1}{\sqrt{2\pi\lambda_i^2\delta^2}} \exp\left\{-\frac{(r - n_i\lambda_i - y_i)^2}{2\lambda_i^2\delta^2}\right\} \\ &\propto \exp\left\{-\frac{(r - n_i\lambda_i - y_i)^2}{2\lambda_i^2\delta^2}\right\} \\ &= \exp\left\{-\frac{1}{2\delta^2}\right\} + \exp\left\{-\left(\frac{r}{\lambda_i} - n_i - \frac{y_i}{\lambda_i}\right)^2\right\} \\ &\propto \exp\left\{-\left(\frac{r}{\lambda_i} - n_i - \frac{y_i}{\lambda_i}\right)^2\right\} \end{aligned} \tag{4.2}$$

Then the joint log-likelihood function with m measurements is:

$$\log\prod_{i=1}^{m} f(y_i|r) = \sum_{i=1}^{m} \log f(y_i|r) \tag{4.3}$$

$$\propto \sum_{i=1}^{m}\left\{-\left(\frac{r}{\lambda_i} - n_i - \frac{y_i}{\lambda_i}\right)^2\right\} \tag{4.4}$$

In vector form, the MLE is then equivalent to

$$(\hat{r}, \hat{\mathbf{N}}) = \arg\min_{\boldsymbol{n}\in\mathbb{Z}^m, r\in\mathbb{Z}} \left\|r\bar{\boldsymbol{\lambda}} - \boldsymbol{n} - \bar{\boldsymbol{y}}\right\| \tag{4.5}$$

where $\|\cdot\|$ is the 2-norm function, $\boldsymbol{n} = [n_1, \cdots, n_m]$, $\bar{\boldsymbol{\lambda}} = \left[\frac{1}{\lambda_1}, \cdots, \frac{1}{\lambda_m}\right]$ and $\bar{\boldsymbol{y}} = \left[\frac{y_1}{\lambda_1}, \cdots, \frac{y_m}{\lambda_m}\right]$.

To solve (4.5), we should estimate $\hat{\mathbf{N}}$ at first. Then the estimate of r, $\hat{r}$, could be ob-

tained via differentiating (4.5) with respect to r:

$$\begin{aligned}\frac{\partial}{\partial r}\left\|r\bar{\boldsymbol{\lambda}}-\boldsymbol{n}-\bar{\boldsymbol{y}}\right\| &= 0\\ \sum_{i=1}^{m} r\left(\bar{\boldsymbol{\lambda}}-\boldsymbol{n}-\bar{\boldsymbol{y}}\right) &= 0\\ \Rightarrow \hat{r} &= (\boldsymbol{n}+\bar{\boldsymbol{y}})\bar{\boldsymbol{\lambda}}^{+}\end{aligned}$$

where $\bar{\boldsymbol{\lambda}}^{+}=(\bar{\boldsymbol{\lambda}}^{T}\bar{\boldsymbol{\lambda}})^{-1}\bar{\boldsymbol{\lambda}}^{T}$ is the Moore-Penrose pseudo-inverse of $\bar{\boldsymbol{\lambda}}$.

Substituting $\hat{r}=(\boldsymbol{n}+\bar{\boldsymbol{y}})\bar{\boldsymbol{\lambda}}^{+}$ into (4.5), we can obtain the estimator of $\mathbf{N}$ as:

$$\begin{aligned}\hat{\mathbf{N}} &= \arg\min_{\boldsymbol{n}\in\mathbb{Z}^m}\left\|(\boldsymbol{n}+\bar{\boldsymbol{y}})\bar{\boldsymbol{\lambda}}^{+}\bar{\boldsymbol{\lambda}}-(\boldsymbol{n}+\bar{\boldsymbol{y}})\right\|\\ &= \arg\min_{\boldsymbol{n}\in\mathbb{Z}^m}\left\|\boldsymbol{n}\mathbf{A}-(-\bar{\boldsymbol{y}}\mathbf{A})\right\|\end{aligned} \tag{4.6}$$

where

$$\mathbf{A}=(\bar{\boldsymbol{\lambda}}^{T}\bar{\boldsymbol{\lambda}})^{-1}-\mathbf{I}_m=\begin{pmatrix}\prod_{i=1}^{m-1}\lambda_i & 0 & \cdots & 0\\ 0 & \prod_{i=1}^{m-1}\lambda_i & \cdots & 0\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \cdots & \prod_{i=1}^{m-1}\lambda_i & 0\\ \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_1} & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_2} & \cdots & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_m}\end{pmatrix} \tag{4.7}$$

and $\mathbf{I}_m$ is m-dimensional identity matrix.

Therefore, it is easy to find that (4.6), as introduced in Section 2.1.3, is actually the closest vector problem in lattice with lattice basis $\mathbf{A}$ and the given point $-\bar{\boldsymbol{y}}\mathbf{A}$. To solve the CVP, the basis could be reduced to a simple one using LLL algorithm as in Algorithm 2.1. Fortunately, we have following theorem to guarantee that the the reduced basis of $\mathbf{A}$ is a diagonal one if some mild conditions are satisfied.

Theorem 4.1. If the wavelengths used $[\lambda_1,\cdots,\lambda_m]$ are co-prime, i.e.

$$\lambda_i \perp \lambda_j \quad i,j=1,\cdots,m \quad \text{and} \quad i\neq j$$

The basis **A** defined in (4.7) could be reduced into **B** with following form:

$$\mathbf{B} = \begin{pmatrix} \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_1} & 0 & \cdots & 0 \\ 0 & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}} & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{4.8}$$

□

Proof. Since the matrix **A** has integer entries, a unique $m \times m$ matrix **T** in the Hermite Normal form (HNF) of **A** with integer entries can be found such that

$$\mathbf{UA} = \mathbf{T} \tag{4.9}$$

where **U** is a unimodular matrix i.e. $|\det(\mathbf{U})| = 1$.

Recall

$$\mathbf{A} = \begin{pmatrix} \prod_{i=1}^{m-1}\lambda_i & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \prod_{i=1}^{m-1}\lambda_i & 0 \\ -\frac{\prod_{i=1}^{m}\lambda_i}{\lambda_1} & \cdots & -\frac{\prod_{i=1}^{m}\lambda_i}{\lambda_{m-1}} & 0 \end{pmatrix} \tag{4.10}$$

and assume that **T** is a HNF of **A** expressed as

$$\mathbf{T} = \begin{pmatrix} \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_1} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}} & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix} \tag{4.11}$$

We wish to verify that the matrix **U** is a unimodular matrix.

Let **U** be a square matrix with entries $u_{p,q}$, $p, q = 1, \cdots, m$ and $\mathbf{UA} = \mathbf{T}$. Multiplying

out the equation $\mathbf{UA} = \mathbf{T}$ for $p = 1$, $q = 1, \cdots, m$, we have

$$\begin{cases} u_{1,1} \prod_{i=1}^{m-1} \lambda_i - u_{1,m} \frac{\prod_{i=1}^{m} \lambda_i}{\lambda_1} = \frac{\prod_{i=1}^{m-1} \lambda_i}{\lambda_1} \\ \vdots \\ u_{1,m-1} \prod_{i=1}^{m-1} \lambda_i - u_{1,m} \frac{\prod_{i=1}^{m} \lambda_i}{\lambda_{m-1}} = 0 \end{cases} \tag{4.12}$$

thus

$$\begin{cases} u_{1,1}\lambda_1 - u_{1,m}\lambda_m = 1 \\ \vdots \\ u_{1,m-1}\lambda_{m-1} - u_{1,m}\lambda_m = 0 \end{cases} \tag{4.13}$$

Since $\lambda_1, \cdots, \lambda_m$ are co-prime, the first column elements of $\mathbf{U}$, i.e., $[u_{1,1}, \cdots, u_{1,m}]^T$ in (4.13) can be solved by the Euclidean algorithm. In a similar manner, all other elements of $\mathbf{U}$ (for $p = 1, \cdots, m$) can be found. So, $\mathbf{U}$ can be expressed using known terms as

$$\mathbf{U} = \begin{pmatrix} \frac{1+u_{1,m}\lambda_m}{\lambda_1} & \frac{u_{1,m}\lambda_m}{\lambda_2} & \cdots & u_{1,m} \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{1+u_{2,m}\lambda_m}{\lambda_2} & \cdots & u_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\prod_{i=1}^{m} \lambda_i}{\lambda_1} & \frac{\prod_{i=1}^{m} \lambda_i}{\lambda_2} & \cdots & \frac{\prod_{i=1}^{m} \lambda_i}{\lambda_m} \end{pmatrix} \tag{4.14}$$

For $\mathbf{U}$ to be unimodular, we require $|\mathbf{U}| = 1$. To show this, we use a property of determinant, which says that if all the elements of a row or column of a matrix are formed by two addends, the determinant of this matrix decomposes in the sum of two determinants. Let consider the first row of $\mathbf{U}$, which can be expressed as

$$\begin{aligned} &\left[\frac{1 + u_{1,m}\lambda_m}{\lambda_1}, \frac{u_{1,m}\lambda_m}{\lambda_2} \cdots, u_{1,m}\right] = \left[\frac{1}{\lambda_1}, 0, \cdots, 0\right] + \\ &\left[\frac{u_{1,m}\lambda_m}{\lambda_1}, \frac{u_{1,m}\lambda_m}{\lambda_2}, \cdots, u_{1,m}\right]. \end{aligned} \tag{4.15}$$

Then $\mathbf{U}$ can be written in the sum of two matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ which are identical except

for the first row, i.e.,

$$\mathbf{U}_1 = \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \dots & 0 \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{1+u_{2,m}\lambda_m}{\lambda_2} & \dots & u_{2,m} \\ \dots & \dots & \dots & \dots \\ \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_1} & \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_2} & \dots & \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_m} \end{pmatrix} \tag{4.16}$$

and

$$\mathbf{U}_2 = \begin{pmatrix} \frac{u_{1,m}\lambda_m}{\lambda_1} & \frac{u_{1,m}\lambda_m}{\lambda_2} & \dots & u_{1,m} \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{1+u_{2,m}\lambda_m}{\lambda_2} & \dots & u_{2,m} \\ \dots & \dots & \dots & \dots \\ \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_1} & \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_2} & \dots & \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_m} \end{pmatrix}, \tag{4.17}$$

Therefore, we have $\det(\mathbf{U}) = \det(\mathbf{U}_1) + \det(\mathbf{U}_2)$.

It is easily to verify that $\det(\mathbf{U}_2) = 0$ because the first row and last row of $\mathbf{U}_2$ are colinear. As a consequence, $\det(\mathbf{U}) = \det(\mathbf{U}_1)$.

Let $\mathbf{U} = \mathbf{U}_1$ and repeat the above operation to obtain that $\det(\mathbf{U})$ and so on. Finally, $\det(\mathbf{U})$ is evaluated as

$$\det(\mathbf{U}) = \det \begin{pmatrix} \frac{1}{\lambda_1} & 0 & \dots & 0 \\ \frac{\lambda_3 u_{2,3}}{\lambda_1} & \frac{1}{\lambda_2} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_1} & \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_2} & \dots & \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_m} \end{pmatrix} \tag{4.18}$$

Thus $\det(\mathbf{U}) = 1$, i.e., $\mathbf{U}$ is an unimodular matrix. This result also justifies that the matrix $\mathbf{T}$ is a HNF of $\mathbf{A}$ based on the relation (4.9). □

Then, if the wavelengths used are coprime, the estimation (4.6) is equivalent to

$$\hat{\mathbf{N}} = \arg\min_{\boldsymbol{n}\in\mathbb{Z}^m} \|\boldsymbol{n}\mathbf{B} - (-\bar{\boldsymbol{y}}\mathbf{A})\| \tag{4.19}$$

where $\mathbf{B}$ has the following simple form according to Theorem 4.8:

$$\mathbf{B} = \begin{pmatrix} \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_1} & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{4.20}$$

As a consequence, the nearest point in $\mathbf{n}^T\mathbf{B}$ to $\bar{\mathbf{y}}^T\mathbf{A}$ can be found by the Babai's nearest plane algorithm [59], which is given in Algorithm 2.2. The algorithm returns the nearest point, denoted by $\mathbf{P}^T = [p_1, \cdots, p_m]$, in $\mathbf{n}^T\mathbf{A}$ to $\bar{\mathbf{y}}^T\mathbf{A}$. Therefore, the estimate of $\mathbf{N}^T$ can be computed as

$$\hat{\mathbf{N}}^T \triangleq \left\{ \mathbf{n}^T \in \mathbb{Z}^m | \mathbf{n}^T\mathbf{A} = \mathbf{P} \right\} \tag{4.21}$$

Since $\mathbf{A}$ is not full rank, $\hat{\mathbf{N}}^T$ cannot be computed directly. From $\mathbf{n}^T\mathbf{A} = \mathbf{P}^T$, we have

$$\begin{cases} n_1 \prod_{i=1}^{m-1}\lambda_i - n_m \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_1} = p_1 \\ \vdots \\ n_{m-1} \prod_{i=1}^{m-1}\lambda_i - n_m \frac{\prod_{i=1}^{m}\lambda_i}{\lambda_{m-1}} = p_{m-1} \end{cases} \tag{4.22}$$

where n_i represents an unknown integer, and p_i is the element of $\mathbf{P}^T$, $i = 1, \cdots, m$.

Equation (4.22) consists of congruence equations which are solvable since all the parameters in the equations are integers. The solution is of the form

$$n_i = \tilde{n}_i + kd_i \tag{4.23}$$

where $\tilde{n}_i$ and d_i can be obtained from (4.22), k is an unknown integer. Equation (4.23) is the general solution of (4.22), where $\tilde{n}_i$ is a special solution, d_i can be viewed as the period of the solution and $k \in \mathbb{Z}$ is the only unknown integer to be determined.

A method based on the computation of the Hermite Normal Form (HNF) of the matrix $\mathbf{A}$ is adopted to solve (4.22). Following from (4.9), we have

$$\mathbf{UA} = \mathbf{B} \tag{4.24}$$

where $\mathbf{U}$ is a unimodular matrix ($\det(\mathbf{U}) = 1$).

Since $\mathbf{UA} = \mathbf{B}$, we can write

$$\mathbf{n}^T\mathbf{A} = (\mathbf{n}^T\mathbf{U}^{-1})(\mathbf{UA}) = (\mathbf{n}^T\mathbf{U}^{-1})\mathbf{B} = \mathbf{P}.$$

Let $(\mathbf{n}^T\mathbf{U}^{-1}) = \mathbf{X}^T = [x_1, \cdots, x_m]$. It is easy to see that $\mathbf{X}^T$ is an integer vector since the elements of both $\mathbf{n}^T$ and $\mathbf{U}$ are all integers. Therefore, $\mathbf{n}^T\mathbf{A} = \mathbf{P}^T$ has integer solutions when $\mathbf{X}^T\mathbf{B} = \mathbf{P}^T$ has an integer solution for $\mathbf{X}^T$ and all the solutions of $\mathbf{n}^T\mathbf{A} = \mathbf{P}^T$ are of the form $\mathbf{n}^T = \mathbf{X}^T\mathbf{U}$, where $\mathbf{X}^T$ can be found from the relationship $\mathbf{X}^T\mathbf{B} = \mathbf{P}^T$. When the set of wavelengths λ_i, $i = 1, \cdots, m$ are all co-prime, according to (4.20) we have

$$\mathbf{X}^T = \left[\frac{p_1\lambda_1}{\prod_{i=1}^{m-1}\lambda_i}, \cdots, \frac{p_{m-1}\lambda_{m-1}}{\prod_{i=1}^{m-1}\lambda_i}, x_m \right].$$

Substitution of $\mathbf{X}^T$ into $\mathbf{n}^T = \mathbf{X}^T\mathbf{U}$, yields the following form for the solution of $\hat{\mathbf{N}}^T$

$$\begin{cases} \hat{N}_1 = \tilde{n}_1 + x_m u_{m,1} \\ \qquad\vdots \\ \hat{N}_m = \tilde{n}_m + x_m u_{m,m} \end{cases} \tag{4.25}$$

where $\tilde{n}_1, \cdots, \tilde{n}_m$ are the special solutions of $\hat{\mathbf{N}}^T$, x_m is an unknown integer and $u_{m,i}$ is the (m,i)th entry of $\mathbf{U}$. In similar vein to (4.23), the equation (4.25) is a general solution of $\hat{\mathbf{N}}^T$ with the period $u_{m,i}$ and x_m is the only unknown integer, which may be determined as follows. Since $0 \le \hat{N}_1 < \frac{\mathrm{LCM}(\lambda_1,\cdots,\lambda_m)}{\lambda_1}$, we have

$$x_m = \mathrm{floor}\left\lfloor \frac{\frac{\mathrm{LCM}(\lambda_1,\cdots,\lambda_m)}{\lambda_1} - \tilde{n}_1}{u_{m,1}} \right\rfloor \tag{4.26}$$

or

$$x_m = \mathrm{ceil}\left\lceil \frac{-\tilde{n}_1}{u_{m,1}} \right\rceil \tag{4.27}$$

where $\mathrm{floor}\lfloor\cdot\rfloor$ and $\mathrm{ceil}\lceil\cdot\rceil$ denote the floor and ceiling functions, respectively. Substitution of the solution of x_m into (4.25) gives the following estimated set of integers $\hat{\mathbf{N}}^T = [\hat{N}_1, \cdots, \hat{N}_m]$.

The above algorithm can be implemented using the pseudo-code shown in Algorithm 4.1. As the Steps 2-4 can be pre-computed in advance, the algorithm is computationally efficient and search free.

Algorithm 4.1: The closed-form algorithm to determine $\mathbf{N}$

```
1  begin
2      Construct basis A by [λ_1, ··· , λ_m], compute reduced basis B of A;
3      Obtain B* by applying the Gram-Schmidt process to B;
4      Find the Hermite normal form decomposition of A, obtain B and U;
5      Obtain the remainders y = [y_1, ··· , y_m]^T and compute ȳ by ȳ = [y_1/λ_1, ··· , y_m/λ_m]^T;
6      Compute the nearest point P^T to ȳ^T A according to Algorithm 2.2;
7      Solve X^T from X^T B = P^T and then obtain the general solution of N̂^T by
       n^T = X^T U;
8      Determine x_m by (4.26) or (4.27) and then substituting it into (4.25).
9      return N̂^T
10 end
```

4.1.2 The Performance Analysis of the Lattice Algorithm and Simulation

In this section, the performance analysis of the lattice algorithm is analyzed based on the assumption of noise distribution.

Recall the estimator (4.6), we have

$$\arg\min_{\boldsymbol{n}\in\mathbb{Z}^m} \|\boldsymbol{n}\mathbf{A} - (-\bar{\boldsymbol{y}}\mathbf{A})\| \tag{4.28}$$

$$= \arg\min_{\boldsymbol{n}\in\mathbb{Z}^m} \|\boldsymbol{n}\mathbf{A} - (-(\bar{\boldsymbol{y}}_{true} + \bar{\omega})\mathbf{A})\| \tag{4.29}$$

$$= \arg\min_{\boldsymbol{n}\in\mathbb{Z}^m} \|\boldsymbol{n}\mathbf{A} - (-\bar{\boldsymbol{y}}_{true}) + \bar{\omega}\mathbf{A}\| \tag{4.30}$$

where $\bar{\omega} \sim \mathcal{N}(0, \delta^2\mathbf{I}_m)$, and $\mathbf{I}_m$ is m-dimensional identity matrix.

According to the CVP, the estimator (4.6) will give the correct estimation of $\mathbf{n}$ if the given point point lies in the Voronoi cell of $(-\bar{\boldsymbol{y}}_{true})$, i.e.

$$(-\bar{\boldsymbol{y}}_{true}) + \bar{\omega}\mathbf{A} \in \boldsymbol{V}(\mathbf{B}, -\bar{\boldsymbol{y}}_{true})$$

where $\boldsymbol{V}(\mathbf{B}, -\bar{\boldsymbol{y}}_{true})$ denotes the Voronoi cell of $-\bar{\boldsymbol{y}}_{true}$ with basis $\mathbf{B}$. On the other hand,

the Voronoi cell is just a copy of that at the origin shifted by a lattice element, so it is sufficient to consider $\bar{\boldsymbol{y}}_{true} = \mathbf{0}$.

Therefore, it is easy to see that the probability of $Pr(\hat{\mathbf{N}} = \mathbf{N}|\delta^2, \lambda_1 \cdots, \lambda_m)$ is

$$Pr(\hat{\mathbf{N}} = \mathbf{N}|\delta^2, \lambda_1 \cdots, \lambda_m) = Pr(\bar{\omega}\mathbf{A} \in \boldsymbol{V}(\mathbf{B}, \mathbf{0})|\delta^2, \lambda_1 \cdots, \lambda_m) \tag{4.31}$$

Since $\bar{\omega} \sim \mathcal{N}(0, \delta^2\mathbf{I}_m)$, then

$$\bar{\omega}\mathbf{A} \sim \mathcal{N}(0, \mathbf{A}^T\delta^2\mathbf{I}_m\mathbf{A}) \tag{4.32}$$

Thus,

$$Pr(\bar{\omega}\mathbf{A} \in \boldsymbol{V}(\mathbf{B}, \mathbf{0})|\delta^2, \lambda_1 \cdots, \lambda_m) \tag{4.33}$$

$$= \int_{\boldsymbol{V}(\mathbf{B},\mathbf{0})} f(\boldsymbol{x})d\boldsymbol{x} \tag{4.34}$$

$$= \int_{-\frac{\Pi_{i=1}^{m-1}\lambda_i}{\lambda_1}}^{\frac{\Pi_{i=1}^{m-1}\lambda_i}{\lambda_1}} \cdots \int_{-\frac{\Pi_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}}}^{\frac{\Pi_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}}} f(x_1, \cdots, x_{m-1})dx_1 \cdots dx_{m-1} \tag{4.35}$$

where

$$f(\boldsymbol{x}) = \frac{1}{2(\pi|\mathbf{A}^T\delta^2\mathbf{I}_m\mathbf{A}|)^{\frac{m-1}{2}}} \exp\left\{-\frac{1}{2}\boldsymbol{x}^T(\mathbf{A}^T\delta^2\mathbf{I}_m\mathbf{A})^{-1}\boldsymbol{x}\right\}$$

An example of the 2-D Voronoi cell of lattice space and measurement error ellipsoid are shown in Fig.4.1, where the ellipsoid $\mathbf{E}$ is defined by the $f(\boldsymbol{x})$ and a confidence level.

4.1.3 Conclusions

In this section, the problem of distance estimation using ambiguous and noisy phase measurements is formulated in the framework of lattice theory and the lattice based algorithm[27] is introduced. It is shown that it has similar performance to the closed-form CRT algorithm[27]. The lattice based algorithm is extremely fast and it computational complexity only depends on the number of wavelengths used.

However, there is a crucial condition on this algorithm, namely the wavelengths used must be co-prime. Without this assumption, the diagonal reduced basis may not exist and

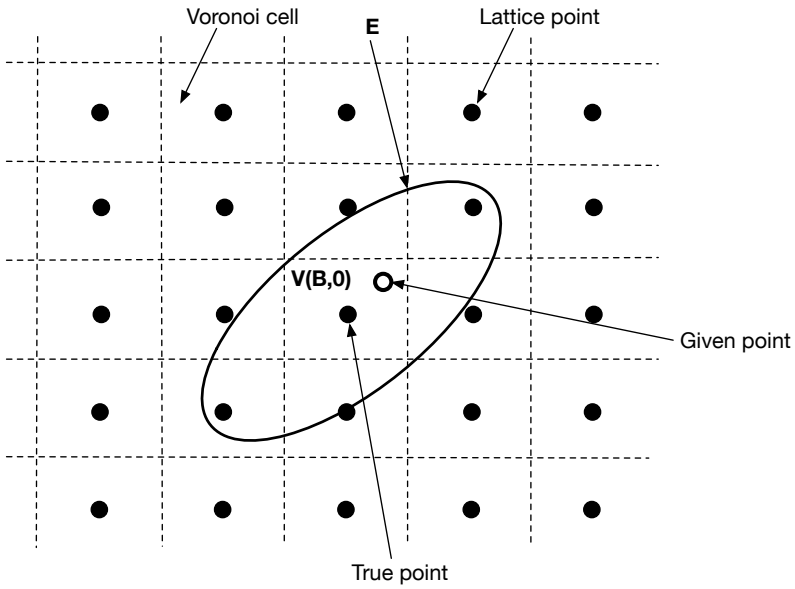


Figure 4.1: The example of Voronoi cell of lattice

the fast closest point searching algorithm is inapplicable. This may limit the applications of the lattice algorithm in practice. Therefore, we are interested in relaxing the co-prime condition.

On the other hand, in estimating distance using ambiguous phase, the estimation of ambiguous integer **N** is particularly crucial since if the $\boldsymbol{n}$ is estimated incorrectly, the distance estimation will be far from the truth. Therefore, to increase the performance of the successful reconstruction of **N** is important in real world.

4.2 A Relaxed Condition on Wavelengths for the Closed-form Algorithm

4.2.1 Introduction

In Section 4.1, the lattice algorithm for solving noisy congruence equations is described and analysed. It should be noted that the efficiency of this algorithm partly derives from the fact that the reduced basis **B** of the original lattice basis is diagonal based in turn on the co-primality assumption on the wavelengths used. Then Babai's algorithm which only needs $m-1$ times computations is applicable and the optimal result(the closest point) can be guaranteed.

The coprimality requirement may prevent the algorithm being used for many real applications as, additionally, the wavelengths used are normally constrained to lie within a limited bandwidth [13, 28]. The latter imposes limitations on the maximum distance that can be unambiguously measured. For example, the admissible wavelengths in the 2.4GHz Wi-fi band are $\{125, 124, \cdots, 120\}$mm, and the coprime wavelengths in this interval may not cover a sufficient large unambiguous distance as shown in the Section 4.2.4. The coprimality condition may also reduce the resolution of range; for example, in a medium-to-high PRF radar, the target may be unobservable because of the so-called "blind zone" and clutter [29] when the number of available PRFs is not enough. Clearly, the practical effectiveness of the algorithm is limited by the coprimality condition.

In this section, we show that the coprimality condition on the wavelengths in the closed-form lattice algorithm can be relaxed to a more general condition that allows a larger number of integer-related wavelengths to be used . As a result, a diagonal reduced basis of $\mathbf{A}$ defined in (4.7) is still guaranteed and a closed-form lattice algorithm is implementable. In addition, the construction of the associated unimodular matrix for the implementation of the closed-form lattice method is derived. The performance of the lattice algorithm under this new condition is analyzed and demonstrated via simulation.

4.2.2 The Proof of the Relaxed Condition

To prove the main result (Theorem 4.2) of this section, the following three propositions are required.

Proposition 4.1. Let $\{a, b, k\} \in \mathbb{Z}^+$. Then

$$\frac{\mathrm{GCD}\,(a, kb)}{\mathrm{GCD}\,(a, b)} \in \mathbb{Z}^+$$

□

Proof. Since there exist $\{u, v\} \in \mathbb{Z}$ such that $\mathrm{GCD}\,(a, kb) = au + kbv$, and $\left\{ \frac{a}{\mathrm{GCD}(a,b)}, \frac{b}{\mathrm{GCD}(a,b)} \right\} \in$

$\mathbb{Z}$, we have

$$\begin{aligned}\frac{\text{GCD}\,(a,kb)}{\text{GCD}\,(a,b)} =& \frac{au+kbv}{\text{GCD}\,(a,b)}\\ =& \frac{a}{\text{GCD}\,(a,b)}u+\frac{b}{\text{GCD}\,(a,b)}kv\\ =& \text{GCD}\left(\frac{a}{\text{GCD}\,(a,b)},k\frac{b}{\text{GCD}\,(a,b)}\right)\in\mathbb{Z}^+\end{aligned}$$ □

Proposition 4.2. Let $d_{i,m}=\text{GCD}\,(\lambda_i,k_i\lambda_m)$ and $g_{i,m}=\text{GCD}(\lambda_i,\lambda_m)$, $i=1,\cdots,m-1$, where $k_i\in\mathbb{Z}^+$. Then

$$\prod_{i=1}^{m-1}d_{i,m}=\prod_{i=1}^{m-1}g_{i,m}$$

if and only if $d_{i,m}=g_{i,m},\ \ i=1,\cdots,m-1$. □

Proof. It is easy to see that

$$d_{i,m}=g_{i,m},\ \ i=1,\cdots,m-1\Longrightarrow\prod_{i=1}^{m-1}d_{i,m}=\prod_{i=1}^{m-1}g_{i,m}$$

On the other hand, from Proposition 4.2, we have

$$\frac{d_{i,m}}{g_{i,m}}=\text{GCD}\left(\frac{\lambda_i}{g_{i,m}},k_i\frac{\lambda_m}{g_{i,m}}\right)\geq 1 \tag{4.36}$$

Then

$$\prod_{i=1}^{m-1}d_{i,m}\geq\prod_{i=1}^{m-1}g_{i,m},$$

with equality holds if $d_{i,m}=g_{i,m}\ \ i=1,\cdots,m-1$. □

Proposition 4.3. Let $\{a,b,c\}\in\mathbb{Z}^+$. Then $\text{GCD}(a,bc)=1$ if and only if $\text{GCD}(a,b)=1$ and $\text{GCD}(a,c)=1$. □

Proof. If $\text{GCD}(a,bc)=1$, since

$$\text{GCD}(a,bc)\geq\text{GCD}(a,b)$$

and

$$\mathrm{GCD}(a, bc) \geq \mathrm{GCD}(a, c),$$

then we have $\mathrm{GCD}(a, b) = 1$ and $\mathrm{GCD}(a, c) = 1$.

It is obvious that if $\mathrm{GCD}(a, b) = 1$ and $\mathrm{GCD}(a, c) = 1$, then $\mathrm{GCD}(a, bc) = 1$. □

Now, we present the main result.

Theorem 4.2. Let $\mathbf{A}$ be a matrix defined by (4.7), $\lambda_i \in \mathbb{Z}^+$ and $g_{i,m} = \mathrm{GCD}(\lambda_i, \lambda_m)$ for $i = 1, \cdots, m-1$. If

$$\mathrm{GCD}\left(\frac{\lambda_i}{g_{i,m}}, \frac{\lambda_j}{g_{j,m}}\right) = 1, i, j = 1, \cdots, m-1, \;\; i \neq j.$$

Then there are integer matrices $\mathbf{B}$ and $\mathbf{U}$ satisfying

$$\mathbf{B} = \begin{pmatrix} g_{1,m}\frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_1} & 0 & \cdots & 0 \\ 0 & g_{2,m}\frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_2} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & g_{m-1,m}\frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}} & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix} \tag{4.37}$$

and $\det|\mathbf{U}| = 1$, such that $\mathbf{UA} = \mathbf{B}$. □

Proof. The proof of Theorem 4.2 is in two parts:

1) proof that there exists an integer matrix $\mathbf{U}$ such that $\mathbf{UA} = \mathbf{B}$;

2) proof of $\det|\mathbf{U}| = 1$.

For the first part, let $u_{i,j}$ be the (i, j)-th entry of matrix $\mathbf{U}$ and let $\mathbf{B}$ be of the form:

$$\mathbf{B} = \begin{pmatrix} b_{1,1} & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & b_{m-1,m-1} & 0 \\ 0 & \ldots & 0 & 0 \end{pmatrix} \tag{4.38}$$

Recall the definition of $\mathbf{A}$ in (4.7), and let $\mathbf{UA} = \mathbf{B}$. Then, for the first row of $\mathbf{B}$, we have

$$\begin{cases} u_{1,1}\prod_{l=1}^{m-1}\lambda_l - u_{1,m}\frac{\prod_{l=1}^{m}\lambda_l}{\lambda_1} = b_{1,1} \\ u_{1,2}\prod_{l=1}^{m-1}\lambda_l - u_{1,m}\frac{\prod_{l=1}^{m}\lambda_l}{\lambda_2} = 0 \\ \quad\vdots \\ u_{1,m-1}\prod_{l=1}^{m-1}\lambda_l - u_{1,m}\frac{\prod_{l=1}^{m}\lambda_l}{\lambda_{m-1}} = 0 \end{cases} \tag{4.39}$$

Let $b_{1,1} = d_{1,m}\frac{\prod_{l=1}^{m-1}\lambda_l}{\lambda_1}$, where $d_{1,m} \in \mathbb{Z}$. Then (4.39) becomes

$$\begin{cases} u_{1,1}\lambda_1 - u_{1,m}\lambda_m = d_{1,m} \\ u_{1,2:(m-1)}\lambda_{2:(m-1)} - u_{1,m}\lambda_m = 0 \\ \quad\vdots \\ u_{1,m-1}\lambda_{m-1} - u_{1,m}\lambda_m = 0 \end{cases} \tag{4.40}$$

Let $c_{1,i} \in \mathbb{Z}$, $i = 1, \cdots, m-1$. Consider the last equation in (4.40); that is,

$$u_{1,m-1}\lambda_{m-1} - u_{1,m}\lambda_m = 0$$

The integer solutions are

$$\begin{cases} u_{1,m-1} = \frac{\lambda_m}{g_{m-1,m}}c_{1,m-1} \\ u_{1,m} = \frac{\lambda_{m-1}}{g_{m-1,m}}c_{1,m-1} \end{cases} \tag{4.41}$$

where $g_{m-1,m} = \mathrm{GCD}(\lambda_{m-1}, \lambda_m)$.

Substituting (4.41) into $u_{1,m-2}\lambda_{m-2} - u_{1,m}\lambda_m = 0$ from Equation (4.40) yields

$$u_{1,m-2}\lambda_{m-2} - \lambda_m\frac{\lambda_{m-1}}{g_{m-1,m}}c_{1,m-1} = 0, \tag{4.42}$$

so that the integer solutions are

$$\begin{cases} u_{1,m-2} = \frac{\lambda_m\lambda_{m-1}}{g_{m-2,m}g_{m-1,m}}c_{1,m-2} \\ c_{1,m-1} = \frac{\lambda_{m-2}}{g_{m-2,m}}c_{1,m-2} \end{cases} \tag{4.43}$$

Now substitution of (4.43) into (4.41) produces

$$u_{1,m} = \frac{\lambda_{m-2}\lambda_{m-1}}{g_{m-2,m}g_{m-1,m}} c_{1,m-2}. \tag{4.44}$$

Similarly, using the last $m-2$ equations in (4.40), we obtain

$$u_{1,m} = \frac{\prod_{l=2}^{m-1}\lambda_l}{\prod_{l=2}^{m-1} g_{l,m}} c_{1,2} \tag{4.45}$$

Finally, substitution of (4.45) into $u_{1,1}\lambda_1 - u_{1,m}\lambda_m = d_{1,m}$ gives

$$u_{1,1}\lambda_1 - \frac{\prod_{l=2}^{m}\lambda_l}{\prod_{l=2}^{m-1} g_{l,m}} c_{1,2} = d_{1,m} \tag{4.46}$$

By Bézout's identity [79], $u_{1,1}$ and $c_{1,2}$ have integral solution if $d_{1,m} = \text{GCD}\left(\lambda_1, \frac{\prod_{l=2}^{m}\lambda_l}{\prod_{l=2}^{m-1} g_{l,m}}\right)$. In turn,

$$b_{1,1} = \text{GCD}\left(\lambda_1, \frac{\prod_{l=2}^{m}\lambda_l}{\prod_{l=2}^{m-1} g_{l,m}}\right) \frac{\prod_{l=1}^{m-1}\lambda_l}{\lambda_1}$$

Based on the above analysis, the elements of the ith row, $i = 1, \cdots, m-1$, of $\mathbf{U}$ are all integers if

$$d_{i,m} = \text{GCD}\left(\lambda_i, \frac{g_{i,m}\prod_{l=1}^{m}\lambda_l}{\lambda_i \prod_{l=1}^{m-1} g_{l,m}}\right) = \text{GCD}\left(\lambda_i, k_i\lambda_m\right)$$

where $k_i = \frac{g_{i,m}\prod_{l=1}^{m-1}\lambda_l}{\lambda_i \prod_{l=1}^{m-1} g_{l,m}} \in \mathbb{Z}^+$. It follows that

$$b_{i,i} = d_{i,m}\left(\prod_{l=1}^{m-1}\lambda_l\right) \Big/ \lambda_i \tag{4.47}$$

Since the elements in the mth row of $\mathbf{B}$ are all zero, the following equations can be constructed from $\mathbf{UA} = \mathbf{B}$

$$\begin{cases} u_{m,1}\prod_{l=1}^{m-1}\lambda_l - u_{m,m}\frac{\prod_{l=1}^{m}\lambda_l}{\lambda_1} = 0 \\ \quad\vdots \\ u_{m,m-1}\prod_{l=1}^{m-1}\lambda_l - u_{m,m}\frac{\prod_{l=1}^{m}\lambda_l}{\lambda_{m-1}} = 0 \end{cases} \tag{4.48}$$

From (4.48), it is not hard to see that $u_{m,j} = \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_j \prod_{l=1}^{m-1} g_{l,m}}$ for $j = 1, \cdots, m$.

The above conclusion indicates that we always can find an integer matrix $\mathbf{U}$ so that $\mathbf{UA} = \mathbf{B}$. If we can show that $\det(\mathbf{U}) = 1$, $\mathbf{B}$ will be the reduced lattice basis of $\mathbf{A}$. Using the above analysis, $\mathbf{U}$ can be rewritten in the following form

$$\mathbf{U} = \begin{pmatrix} \frac{d_{1,m}+u_{1,m}\lambda_m}{\lambda_1} & \frac{u_{1,m}\lambda_m}{\lambda_2} & \cdots & u_{1,m} \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{d_{2,m}+u_{2,m}\lambda_m}{\lambda_2} & \cdots & u_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_1 \prod_{l=1}^{m-1} g_{l,m}} & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_2 \prod_{l=1}^{m-1} g_{l,m}} & \cdots & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_m \prod_{l=1}^{m-1} g_{l,m}} \end{pmatrix} \tag{4.49}$$

$\det(\mathbf{U})$ can then be rewritten as $\det(\mathbf{U}) = \det(\mathbf{U}_1) + \det(\mathbf{U}_2)$, where

$$\mathbf{U}_1 = \begin{pmatrix} \frac{d_{1,m}}{\lambda_1} & 0 & \cdots & 0 \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{d_{2,m}+u_{2,m}\lambda_m}{\lambda_2} & \cdots & u_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_1 \prod_{l=1}^{m-1} g_{l,m}} & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_2 \prod_{l=1}^{m-1} g_{l,m}} & \cdots & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_m \prod_{l=1}^{m-1} g_{l,m}} \end{pmatrix}$$

and

$$\mathbf{U}_2 = \begin{pmatrix} \frac{u_{1,m}\lambda_m}{\lambda_1} & \frac{u_{1,m}\lambda_m}{\lambda_2} & \cdots & u_{1,m} \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{d_{2,m}+u_{2,m}\lambda_m}{\lambda_2} & \cdots & u_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_1 \prod_{l=1}^{m-1} g_{l,m}} & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_2 \prod_{l=1}^{m-1} g_{l,m}} & \cdots & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_m \prod_{l=1}^{m-1} g_{l,m}} \end{pmatrix}$$

As the first and last row in $\mathbf{U}_2$ are collinear, then

$$\det(\mathbf{U}) = \det \begin{bmatrix} \frac{d_{1,m}}{\lambda_1} & 0 & \cdots & 0 \\ \frac{u_{2,m}\lambda_m}{\lambda_1} & \frac{d_{2,m}+u_{2,m}\lambda_m}{\lambda_2} & \cdots & u_{2,m} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_1 \prod_{l=1}^{m-1} g_{l,m}} & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_2 \prod_{l=1}^{m-1} g_{l,m}} & \cdots & \frac{\prod_{l=1}^{m}\lambda_l}{\lambda_m \prod_{l=1}^{m-1} g_{l,m}} \end{bmatrix} \tag{4.50}$$

Similarly, we have following

$$\det(\mathbf{U}) = \left(\prod_{i=1}^{m-1} \frac{d_{i,m}}{\lambda_i}\right) \frac{\prod_{l=1}^{m} \lambda_l}{\lambda_m \prod_{l=1}^{m-1} g_{l,m}} = \frac{\prod_{l=1}^{m-1} d_{l,m}}{\prod_{l=1}^{m-1} g_{l,m}} \tag{4.51}$$

If $\prod_{l=1}^{m-1} d_{l,m} = \prod_{l=1}^{m-1} g_{l,m}$, then $\det(\mathbf{U}) = 1$. From Proposition 4.2, then, $\prod_{l=1}^{m-1} d_{l,m} = \prod_{l=1}^{m-1} g_{l,m}$ if and only if $g_{i,m} = d_{i,m}, i = 1, \cdots, m-1$.

In other words, $\mathbf{U}$ is a unimodular matrix and, in view of (4.47) and (4.38), $\mathbf{B}$ has the form

$$\mathbf{B} = \begin{pmatrix} b_{1,1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & b_{m-1,m-1} & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix} \tag{4.52}$$

if and only if $g_{i,m} = d_{i,m}, i = 1, \cdots, m-1$, where $b_{i,i} = g_{i,m} \frac{\prod_{l=1}^{m-1} \lambda_l}{\lambda_i}$.

In particular, from Proposition 4.3, we remark that the condition for $\mathbf{B}$ to be of the form (4.52) is $(i, j = 1, \cdots, m-1, i \neq j)$

$$\begin{aligned} 1 = \frac{d_{i,m}}{g_{i,m}} &= \mathrm{GCD}\left(\frac{\lambda_i}{g_{i,m}}, \left(\frac{\prod_{l=1}^{m-1} \lambda_l}{\lambda_i} \Big/ \frac{\prod_{l=1}^{m-1} g_l}{g_i}\right) \frac{\lambda_m}{g_{i,m}}\right) \\ &\Leftrightarrow \mathrm{GCD}\left(\frac{\lambda_i}{g_{i,m}}, \frac{\lambda_j}{g_{j,m}}\right) = 1 \end{aligned}$$ □

Theorem 4.2 implies that the matrix $\mathbf{B}$ defined in (4.37) is a reduced lattice basis for $\mathbf{A}$. The following lemma is trivial. It states that the coprimality condition in [27] is a special case of Theorem 4.2.

Lemma 4.1. If $\mathrm{GCD}(\lambda_i, \lambda_j) = 1, \quad i, j = 1, \cdots, m, i \neq j$, then $\mathrm{GCD}\left(\frac{\lambda_i}{g_{i,m}}, \frac{\lambda_j}{g_{j,m}}\right) = 1, i, j = 1, \cdots, m-1, \quad i \neq j$. □

As shown in [27], the associated unimodular matrix $\mathbf{U}$ is essential for implementing the closed-form lattice algorithm. Now we show how to determine the values of $\mathbf{U}$. In view of (4.49), it is sufficient to compute $u_{i,m}$ and $d_{i,m}, i = 1, \cdots, m-1$ so as to determine

U. From the relation

$$u_{i,i}\lambda_i - \frac{g_{i,m}\prod_{l=1}^{m}\lambda_l}{\lambda_i\prod_{l=1}^{m-1}g_{l,m}}c_{i,2} = d_{i,m} \tag{4.53}$$

and $d_{i,m} = g_{i,m}$, we have

$$u_{i,i}\frac{\lambda_i}{g_{i,m}} - \left(\frac{\prod_{l=1}^{m-1}\lambda_l}{\lambda_i}\Big/\frac{\prod_{l=1}^{m-1}g_l}{g_i}\right)\frac{\lambda_m}{g_{i,m}}c_{i,2} = 1 \tag{4.54}$$

Since GCD $\left(\frac{\lambda_i}{g_{i,m}}, \frac{\lambda_j}{g_{j,m}}\right) = 1,\ i,j = 1,\cdots,m-1, i \neq j$, Equation (4.54) is solvable using the Extended Euclidean algorithm [80]. Consequently, $u_{i,m}$ is determined by

$$u_{i,m} = \frac{g_{i,m}\prod_{l=1}^{m-1}\lambda_l}{\lambda_i\prod_{l=1}^{m-1}g_{l,m}}c_{i,2} \tag{4.55}$$

and then the unimodular matrix **U** is constructed.

4.2.3 The Performance Analysis of the Lattice Algorithm Using the Relaxed Condition

In this section, the performance of the closed-form lattice algorithm [27] in terms of the reconstruction probability under the relaxed condition is derived using the properties of the lattice. We assume that the signal wavelengths satisfy the condition in Theorem 4.2.

Recall the definition of **A**, since $\bar{\mathbf{y}} = \bar{\mathbf{y}}_{true} + \bar{\omega}$, where

$$\bar{\omega} = \left[\frac{\omega_1}{\lambda_1}, \cdots, \frac{\omega_m}{\lambda_m}\right], \quad \bar{\mathbf{y}}_{true} = \left[\frac{y_{1,true}}{\lambda_1}, \cdots, \frac{y_{m,true}}{\lambda_m}\right]$$

and $\omega_i \sim \mathcal{N}(0, \lambda_i^2\delta^2)$, then

$$\bar{\mathbf{y}}\mathbf{A} = \bar{\mathbf{y}}_{true}\mathbf{A} + \bar{\omega}\mathbf{A} \text{ and } \bar{\omega}\mathbf{A} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}) \tag{4.56}$$

and $\mathbf{\Sigma} = \delta^2 \left(\prod_{i=1}^{m} \lambda_i\right)^2 \mathbf{\Sigma}'$, where $\sigma'_{i,j}$ is the (i,j)-th entry of $\mathbf{\Sigma}'$ and

$$\begin{cases} \sigma'_{i,i} = & \frac{1}{\lambda_i^2} + \frac{1}{\lambda_m^2}, \quad i,j = 1,\cdots,m-1 \text{ and } i \neq j \\ \sigma'_{i,j} = & \frac{1}{\lambda_i \lambda_j}, \qquad i = 1,\cdots,m-1 \end{cases}$$

On the assumption that the condition in Theorem 4.2 is satisfied, the closed-form lattice algorithm will return the correct value of $\mathbf{N}$ if $-(\bar{\mathbf{y}}_{true} + \bar{\omega})\mathbf{A}$ lies in the Voronoi cell of $\mathbf{NB}$, where $\mathbf{B}$ is a diagonal matrix. The Voronoi cell of the diagonal basis $\mathbf{B}$ is an $(m-1)$-dimensional rectangle with vertices are given by

$$\begin{aligned} &\left\{ -\frac{b_{1,1}}{2} + v_1 \frac{b_{1,1}}{2}, \frac{b_{1,1}}{2} + v_1 \frac{b_{1,1}}{2} \right\} \times \cdots \times \\ &\left\{ -\frac{b_{m-1,m-1}}{2} + v_{m-1} \frac{b_{m-1,m-1}}{2}, \frac{b_{m-1,m-1}}{2} + v_{m-1} \frac{b_{m-1,m-1}}{2} \right\} \end{aligned}$$

where $\times$ is the Cartesian product, $\{v_1,\cdots,v_{m-1}\} \in \{0, \pm 1, \pm 3, \pm 5 \cdots\}$ and $b_{i,i} = g_{i,m} \frac{\prod_{i=l}^{m-1} \lambda_l}{\lambda_i}$.

Because of the translational symmetry of the lattice, the Voronoi cells of all lattice points are congruent, [58], we may simply discuss the case $\bar{\mathbf{y}}_{true} = \mathbf{0}$; that is, $\mathbf{N} = \mathbf{0}$. Based on (4.56), we see that the probability of correctly estimating $\mathbf{N}$ given different values of δ and wavelengths is

$$\begin{aligned} &Pr(\hat{\mathbf{N}} = \mathbf{N} | \delta, \lambda_1, \cdots, \lambda_m) \\ =&Pr(\hat{\mathbf{N}} = \mathbf{N} = \mathbf{0} | \delta, \lambda_1, \cdots, \lambda_m) \\ =&\int_{-\frac{b_{1,1}}{2}}^{\frac{b_{1,1}}{2}} \cdots \int_{-\frac{b_{m-1,m-1}}{2}}^{\frac{b_{m-1,m-1}}{2}} \frac{1}{\sqrt{2\pi}|\mathbf{\Sigma}|^{1/2}} \exp\left\{ -\frac{1}{2} \boldsymbol{x} \mathbf{\Sigma}^{-1} \mathbf{x}^T \right\} d\mathbf{x} \end{aligned}$$

where $\mathbf{x} = [x_1,\cdots,x_{m-1}]$.

4.2.4 Simulation

In this section, we demonstrate the distance estimation performance improvement when using the closed-form lattice algorithm with the relaxed hypothesis in Theorem 4.2. We consider a 2.4GHz Wi-fi band transmitter localisation example. Typically, the localisation of a transmitter can be done via triangulation by a receiver able to measure the phase

difference of Wi-fi signals transmitted from the transmitter at multiple frequencies. We assume that the measurement errors are zero-mean Gaussian distributed with standard deviations proportional to their corresponding wavelengths (i.e., $\propto \delta\lambda_i$, $0 < \delta \leq 1$).

Within the 2.4GHz Wi-fi band, the available integer wavelengths are the set (2.40GHz–2.49GHz)

$$\mathbf{\Lambda}_1 = \{125, 124, 123, 122, 121, 120\}\text{mm}.$$

Since $\mathbf{\Lambda}_1$ satisfies the condition in Theorem 4.2, we can use all of them to calculate the distance between the transmitter and receiver. On the other hand, if we assume that the wavelengths satisfying the coprimality condition, as in [27], are used, we would have the following two wavelength sets:

$$\mathbf{\Lambda}_2 = \{125, 124, 123, 121\}\text{mm}$$
$$\mathbf{\Lambda}_3 = \{121, 120\}\text{mm}$$

Clearly, $\text{GCD}(\mathbf{\Lambda}_3) = 121 \times 120 = 14.5\text{m}$ covers too small an unambiguous distance to be used in the application. We chose $\mathbf{\Lambda}_2$ for the wavelengths to be used under the coprimality condition for "measuring" the distance using the closed-form lattice algorithm. Therefore, we compared the performance of the two closed-form algorithms implemented using $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$, respectively.

Figure 4.2 shows the performance comparison of estimating the unwrapped signal phase (distance) versus different error level δ. The performance is measured in terms of the probability that the correct $\mathbf{N}$ is found.

In the simulation, the distance r (corresponding to the ground truth signal phase) is randomly selected between 1 and the LCM of the wavelength used. The "noise" parameter δ in the phase measurement noise variance $\delta^2\lambda_i^2$ is chosen to be $-20\log_{10}\delta = 20 : 5 : 80$. All simulation results were averaged over 5000 Monte Carlo runs. The theoretic probabilities; that is, $Pr(\hat{\mathbf{N}} = \mathbf{N}|\delta, \mathbf{\Lambda}_{1,2})$, are also plotted in Figure 4.2.

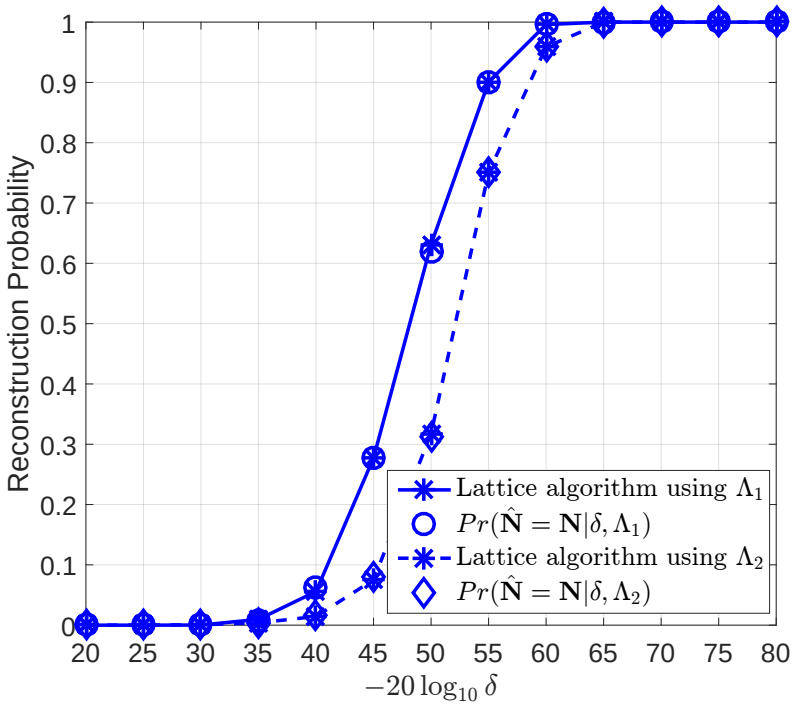


Figure 4.2: Reconstruction probability versus error level

4.2.5 Conclusions

In this section, the standard coprimality condition on waveform sets for estimation of travel distances of wireless signal using a closed-form lattice based algorithm is relaxed. According to this new result, more signal wavelengths may be selected to unwrap signal phase using phase measurements under bandwidth constraints using the closed-form lattice algorithm [27]. The associated probabilistic performance measure and an illustrative simulation example are also given.

4.3 An Algorithm for Phase Unwrapping Using Non-coprime Wavelengths

4.3.1 Introduction

In Section 4.2, a relaxed condition on the closed-form lattice algorithm is presented using the properties of lattice. However, in some practical cases, the available wavelengths

cannot satisfy the new condition. Without the co-prime constraint and new condition, an optimal solution may be achieved via the search-based algorithm [81] or lattice algorithm using the sphere decoder as described in Algorithm 2.3. However, the implementation demands very high computational complexity because of the involvement of exhaustive searches over parameter space.

In this section, we present a lattice based algorithm to solve this general ambiguity problem without the co-prime constraint. Based on the work in [31], a modified iteration for closest point searching is derived so that the proposed algorithm is computationally more efficient than existing approaches. The efficiency of the proposed algorithm is demonstrated via simulation.

4.3.2 The Proposed Algorithm

In this section, we are interested in using moduli which do not satisfy the condition of Theorem 4.2 and therefore, the closed-form lattice algorithm is not applicable.

Recall

$$\mathbf{A} = \begin{pmatrix} \prod_{i=1}^{m-1} \lambda_i & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \prod_{i=1}^{m-1} \lambda_i & 0 \\ -\frac{\prod_{i=1}^{m} \lambda_i}{\lambda_1} & \dots & -\frac{\prod_{i=1}^{m} \lambda_i}{\lambda_{m-1}} & 0 \end{pmatrix} \tag{4.57}$$

By the Hermite Normal Form(HNF) decomposition, $\mathbf{A}$ can be represented as the product of an upper-triangular matrix $\mathbf{T}$ and a unimodular matrix $\mathbf{U}$ [80], so that $\mathbf{\Omega}(\mathbf{A}) = \mathbf{\Omega}(\mathbf{T})$, where $\mathbf{\Omega}(\cdot)$ denotes the lattice space generated by the corresponding basis. The elements of the mth row and column of $\mathbf{T}$ are all zero since the rank of $\mathbf{A}$ is $m-1$.

Suppose that the closest lattice point to the given point $\bar{\mathbf{y}}\mathbf{A}$ and the associated integer vector corresponding to $\mathbf{\Omega}(\mathbf{T})$ are solved and denoted by $\mathbf{P} = [P_1, \cdots, P_{m-1}, 0]$ and $\mathbf{v} = [v_1, \cdots, v_{m-1}, v_m]$ respectively, where v_m is undetermined. Then we have $\mathbf{vT} = \hat{\mathbf{N}}\mathbf{A} = \mathbf{P}$. We aim to determine $\hat{\mathbf{N}}$ from these known values. Since $\mathbf{vT} = (\hat{\mathbf{N}}\mathbf{U}^{-1})\mathbf{T}$, therefore $\mathbf{vU} =$

$\hat{\mathbf{N}}$ which can be written as follows, where $j = 1, \cdots, m$

$$\sum_{i=1}^{m-1} v_i u_{i,j} + v_m u_{m,j} = \hat{N}_j \tag{4.58}$$

All the values are known in these equations except $\hat{N}_j$ and the undetermined integer value v_m . $\hat{N}_j$ in (4.58) are integers and bounded by $0 < \hat{N}_j \leq \frac{\text{LCM}(\lambda_1, \cdots, \lambda_m)}{\lambda_j}$, therefore

$$0 < v_m + \frac{\sum_{i=1}^{m-1} v_i u_{i,j}}{u_{m,j}} \leq \frac{\text{LCM}(\lambda_1, \cdots, \lambda_m)}{\lambda_j u_{m,j}} \tag{4.59}$$

Lemma 4.2. Let $\mathbf{U}$, $\mathbf{A}$ and $\mathbf{T}$ be integer matrices as defined in this section. Then there exist $\{u_{m,j} \in \mathbb{Z}, j = 1, \cdots, m\}$ satisfying following relation

$$\frac{1}{u_{m,1}\lambda_1} = \cdots = \frac{1}{u_{m,m}\lambda_m} = \pm\frac{1}{\text{LCM}(\lambda_1, \cdots, \lambda_m)}$$

□

Proof. Since $\mathbf{UA} = \mathbf{T}$ and the entries of last row of $\mathbf{T}$ are all 0, then we have $u_{m,j}\prod_{i=1}^{m-1}\lambda_i - u_{m,m}\frac{\prod_{i=1}^{m}\lambda_i}{\lambda_j} = 0$ for $j = 1, \cdots, m$

Clearly, the solution of $\{u_{m,j}\}$ is $\frac{K}{\lambda_j}$. Consider $\{u_{m,j} \in \mathbb{Z}, j = 1, \cdots, m\}$, then let $K = k\text{LCM}(\lambda_1, \cdots, \lambda_m)$ where $k \in \mathbb{Z} \setminus 0$ ($\det \mathbf{U}$ will be 0 if $k = 0$). Substituting K into $u_{m,j} = \frac{K}{\lambda_j}$, we have

$$\frac{1}{u_{m,1}\lambda_1} = \cdots = \frac{1}{u_{m,m}\lambda_m} = \pm\frac{1}{k\text{LCM}(\lambda_1, \cdots, \lambda_m)}$$

Let $k = \pm 1$ and the lemma is proved. □

Let $u_{m,j} \in \mathbb{Z}^+$ and $j = 1$, from Lemma 4.2, (4.59) can be written as

$$-\frac{\sum_{i=1}^{m-1} v_i u_{i,1}}{u_{m,1}} < v_m \leq 1 - \frac{\sum_{i=1}^{m-1} v_i u_{i,1}}{u_{m,1}}$$

Lemma 4.3. Let $X \in \mathbb{R}$. there exists an unique integer v_m satisfies $-X < v_m \leq 1 - X$. □

Since $v_m \in \mathbb{Z}$, it follows from Lemma 4.3 that v_m can be uniquely determined

$$\begin{cases} v_m = -\left\lfloor \frac{\sum_{i=1}^{m-1} v_i u_{i,1}}{u_{m,i}} \right\rfloor & \frac{\sum_{i=1}^{m-1} v_i u_{i,1}}{u_{m,i}} \notin \mathbb{Z} \\ v_m = 1 - \frac{\sum_{i=1}^{m-1} v_i u_{i,1}}{u_{m,i}} & \frac{\sum_{i=1}^{m-1} v_i u_{i,1}}{u_{m,i}} \in \mathbb{Z} \end{cases}$$

Therefore, all parameters in (4.58) are known, and $\hat{\mathbf{N}}$ can be uniquely determined.

4.3.3 Finding the Closest Lattice Point By Relevant Vectors

Finding the lattice point in $\mathbf{\Omega}(\mathbf{G})$, where $\mathbf{G}$ is a lattice basis, which is closest to a given point $\mathbf{x} \in \mathbb{R}^m$ can be iteratively implemented as in [31]:

$$\mathbf{t}_{k+1} = \mathbf{t}_k + \mathbf{d}_k \tag{4.60}$$

$$\mathbf{d}_k = \arg \min_{\mathbf{d} \in \mathrm{Rel}(\mathbf{G}) \cup \{\mathbf{0}\}} \|\mathbf{x} - \mathbf{t}_k - \mathbf{d}\| \tag{4.61}$$

with an initial guess on $\mathbf{t}_0$ which lies in $\mathbf{\Omega}(\mathbf{G})$. It can be proved that this algorithm converges to the closest point of $\mathbf{x}$ within a finite number of steps.

Inspired by [31] and [63], we may use an adaptive step size α_k rather a fixed step size in the iteration. Thus, the above iteration may be written as

$$\mathbf{t}_{k+1} = \mathbf{t}_k + \alpha_k \mathbf{d}_k \tag{4.62}$$

$$\{\mathbf{d}_k, \alpha_k\} = \arg \min_{\mathbf{c} \in \mathrm{Rel}(\mathbf{G}),\, \alpha \in \mathbb{Z}^+ \cup \{\mathbf{0}\}} \|\mathbf{x} - \mathbf{t}_k - \alpha \mathbf{c}\| \tag{4.63}$$

We have following proposition to optimally choose α_k:

Proposition 4.4. The optimal choice of α_k satisfies $\alpha_k = Round\left(\left| \frac{\sum_{i=1}^{m-1} \beta_i c_i}{\sum_{i=1}^{m-1} c_i^2} \right| \right)$, where c_i and β_i is ith element of $\mathbf{c}$ and $\mathbf{x} - \mathbf{t}_k$ respectively and $\mathbf{c} \in \mathrm{Rel}(\mathbf{G})$. □

Proof. Given $\mathbf{x}$ and $\mathbf{c}$, $\min_{\alpha \in \mathbb{R}} \|\mathbf{x} - \mathbf{t}_k - \alpha \mathbf{c}\|$ can be written into $\min_\alpha \sum_{i=1}^{m-1} (\beta_i - \alpha c_i)^2$

Taking the derivative of $\sum_{i=1}^{m-1} (\beta_i - \alpha c_i)^2$ with respect to α and setting it zero gives

$$\alpha = \frac{\sum_{i=1}^{m-1} \beta_i c_i}{\sum_{i=1}^{m-1} c_i^2}.$$

Since α_k is either a positive integer or 0, the optimal choice is

$$\alpha_k = Round\left(\left|\frac{\sum_{i=1}^{m-1}\beta_i c_i}{\sum_{i=1}^{m-1}c_i^2}\right|\right).$$

□

Proposition 4.5. The algorithm (4.62) will be converges to the closest point in a finite steps. □

Proof. Denote the closest point of $\mathbf{x}$ by $\mathbf{P}$. If $\mathbf{t}_k$ lies in the Voronoi cell of $\mathbf{P}$, then from the definition of Voronoi region, we have $\sum_{i=1}^{m-1}\beta_i c_i < \frac{1}{2}\sum_{i=1}^{m-1}c_i^2$, where c_i and β_i is ith element of $\mathbf{c}$ and $\mathbf{x}-\mathbf{t}_k$, $\forall \mathbf{c} \in \mathrm{Rel}(\mathbf{G})$, thus from Proposition 4.4, $\alpha_k = 0$.

If $\mathbf{t}_k$ does not lie in the Voronoi cell of $\mathbf{P}$, then from $\mathbf{t}_{k+1} = \mathbf{t}_k + \alpha_k \mathbf{d}_k$, we have

$$\|\mathbf{x}-\mathbf{t}_{k+1}\| \begin{cases} < \|\mathbf{x}-\mathbf{t}_k-\alpha_k\mathbf{c}\| & \alpha \neq 0 \\ = \|\mathbf{x}-\mathbf{t}_{k+1}\| & \alpha_k = 0 \end{cases}$$

where $\alpha \in \mathbb{Z}^+$,$\alpha \neq \alpha_k$, $\forall \mathbf{c} \in \mathrm{Rel}(\mathbf{G})$ and $\mathbf{d}_k \neq \mathbf{c}$.

This implies that the distance between $\mathbf{x}$ and $\mathbf{t}_{k+1}$ will be strictly decreasing until $\alpha_k = 0$. From the above analysis, $\alpha_k = 0$ means that the $\mathbf{t}_k$ lies in the the Voronoi cell of $\mathbf{P}$ and the closest point is attained. □

An illustrative example showing the search path and number of iteration differences between the conventional searching algorithm and the proposed adaptive searching algorithm is given in Fig. 4.3.

After obtaining the closest point to the given point $\mathbf{P}$, it is easy to find the associated vector $\mathbf{v}$ corresponding to lattice $\mathbf{T}$ via $\mathbf{vT} = \mathbf{P}$ since $\mathbf{T}$ is upper-triangular matrix.

4.3.4 Simulation

In this section, we compare the performance of the proposed algorithm with the search-based CRT algorithm presented in [81]. The latter is regarded as an optimal algorithm.

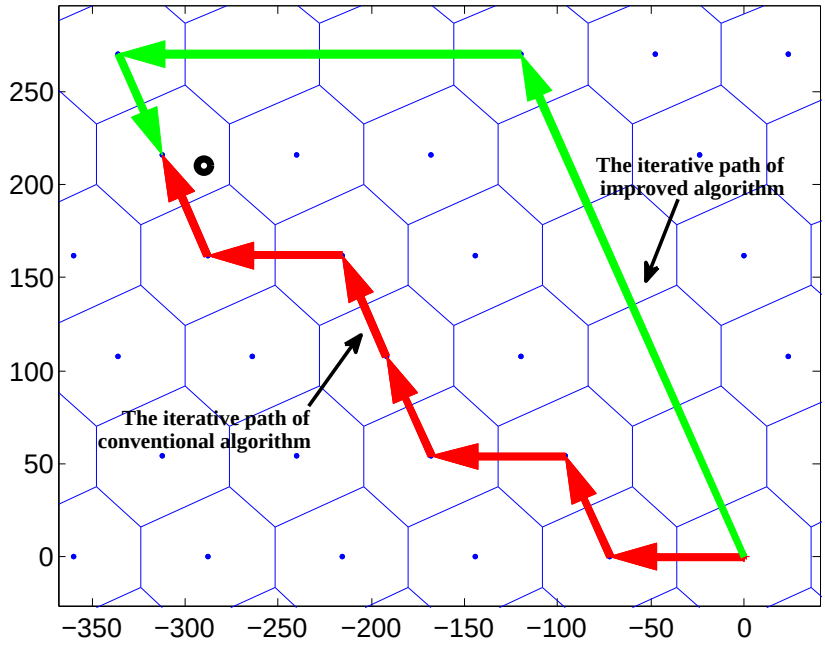


Figure 4.3: An illustrative example for comparing the conventional searching and adaptive searching algorithms. The given point is indicated using a black circle and the lattice point is indicated by blue dot. Red arrowed-line shows the iteration path of conventional algorithm and the green arrowed-line shows the adaptive searching algorithm.

Two sets of wavelengths(moduli), which do not satisfy the co-prime constraint, are used as below to demonstrate the efficiency of the proposed algorithm.

$$\mathbf{\Lambda}_1 = \{21, 22, 23, 24, 25, 26, 27, 29\}$$
$$\mathbf{\Lambda}_2 = \{56, 57, 58, 59, 60, 61, 62, 63\}$$

The distance r is randomly selected between 1 and the LCM of the modulus set. The parameter δ in phase measurement noise variance $\delta^2\lambda_i^2$ is chosen such that $-20\log_{10}\delta = 30 : 2 : 56$, which provides an indication for both noise level and signal to noise ratio in the simulation. All simulation results illustrated are averaged over 1000 Monte Carlo runs. Algorithm performance is measured in the probability of correctly estimating the set of integers $\mathbf{N}$ for a given measurement noise level, i.e., the probability of a correct signal phase reconstruction. Computational complexity is also an important criterion.

Fig.4.4 shows the probabilities of correctly reconstructing signal phases for different

measurement noise levels δ and different moduli sets $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$, i.e. $Pr(\hat{\mathbf{N}} = \mathbf{N}|\delta)$. In both cases ($\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$), the proposed algorithm has an identical performance to the search-based CRT algorithm in the probabilities of correctly estimating both $\mathbf{N}$ and r. On the other hand, Fig.4.5 shows the ratio of required CPU time. It indicates that the required computational load by the proposed algorithm is significantly less than the search-based algorithm[27].

The closest point searching algorithm with an adaptive iteration step size plays a key role for the efficiency enhancement of the proposed algorithm. We compare the computational complexity of the new derived searching algorithm with that of the conventional algorithm presented in [31] in terms of CPU time in Fig. 4.6 versus the number of signal wavelengths used. It shows that the conventional searching algorithm increases its computational overhead much faster than the proposed one once the number of signal wavelengths exceed 8.

It is worth mentioning that if the underlying problem satisfies the co-prime constraint, the proposed algorithm is equivalent to the closed-form lattice algorithm presented in [27] or Section 4.1.

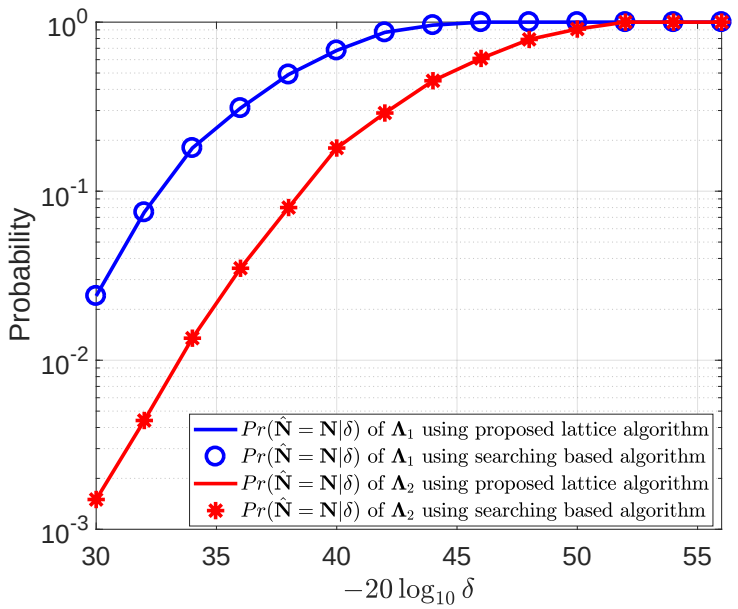


Figure 4.4: Comparison of the proposed algorithm and searching based algorithm vs. the amplitude of the noise using $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$: Probability of correctly reconstruction.

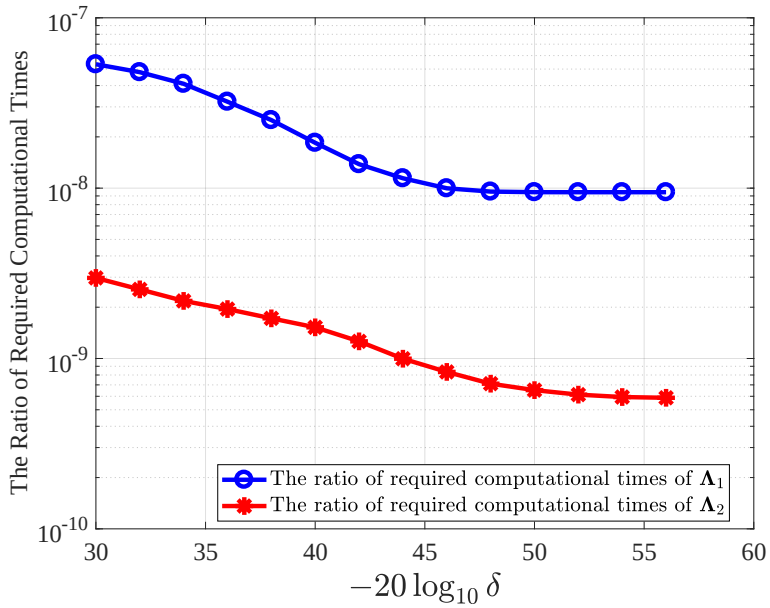


Figure 4.5: Comparison of the proposed algorithm and searching based algorithm vs. the amplitude of the noise using $\mathbf{\Lambda}_1$ and $\mathbf{\Lambda}_2$: Ratio of required computation time.

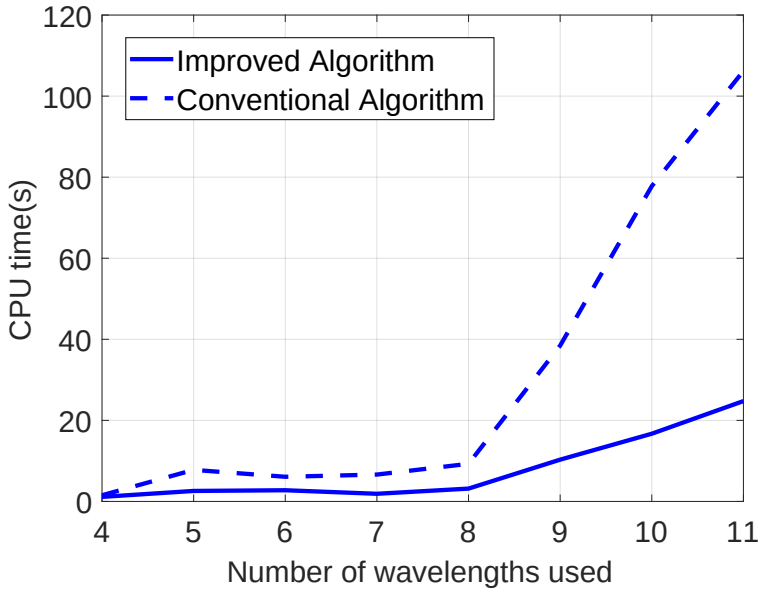


Figure 4.6: The required CPU time of proposed closest point searching algorithm and conventional one versus different number of moduli used.

4.3.5 Conclusions

In this section, we present a lattice based estimator for estimating distances with phase wrapped signal measurements. The proposed algorithm addresses a more general situation where the co-prime constraint on signal wavelengths is relaxed. Furthermore, we propose an adaptive searching algorithm for finding the nearest lattice point for a given noisy measurement on a lattice, which greatly improves the efficiency of the proposed estimator.

4.4 An Improved Lattice Algorithm Using Different Basis Structures

4.4.1 Introduction

In the lattice based approach, the problem is formulated as a state estimation problem, requiring simultaneous estimation of the wrapped signal phase and the unknown integer for the number of times that the range is folded by the signal wavelength. Practically, if the unknown integer is estimated incorrectly, the range estimation error is significant. As a result, the reconstruction probability of the wrapped integer becomes the key measure of algorithm capability for resolving ambiguity.

In this section, we propose an improved lattice based algorithm for localization problem. We show analytically that the proposed algorithm has better performance in terms of reconstruction probability than the one in [27], though with a slightly higher computational complexity.

4.4.2 Improvement Over the Conventional Lattice Algorithm

Before introducing the proposed algorithm, we define two matrices:

$$\mathbf{A}(x_1,\cdots,x_m)=\begin{pmatrix} \prod_{i=1}^{m-1}x_i & \dots & 0 & 0\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \dots & \prod_{i=1}^{m-1}x_i & 0\\ -\frac{\prod_{i=1}^{m}x_i}{x_1} & \dots & -\frac{\prod_{i=1}^{m}x_i}{x_{m-1}} & 0 \end{pmatrix} \tag{4.64}$$

and,

$$\mathbf{B}(x_1,\cdots,x_m)=\begin{pmatrix} \frac{\prod_{i=1}^{m-1}x_i}{x_1} & \dots & 0 & 0\\ \vdots & \ddots & \vdots & \vdots\\ 0 & \dots & \frac{\prod_{i=1}^{m-1}x_i}{x_{m-1}} & 0\\ 0 & \dots & 0 & 0 \end{pmatrix} \tag{4.65}$$

Obviously, $\mathbf{A}(x_1,\cdots,x_m)$ and $\mathbf{B}(x_1,\cdots,x_m)$ are functions of $(x_1,\cdots,x_m)$.

Then recall the following estimator (4.19):

$$\hat{\mathbf{N}}=\arg\min_{\mathbf{n}\in\mathbb{Z}^m}\left\|\mathbf{n}\mathbf{B}(\lambda_1,\cdots,\lambda_m)+\bar{\mathbf{y}}\mathbf{A}(\lambda_1,\cdots,\lambda_m)\right\|. \tag{4.66}$$

Formula (4.66) is typically the closest point searching problem in lattices and can be solved via Babai's algorithm efficiently [27, 60] if the wavelengths are co-prime. This algorithm is referred as the conventional lattice algorithm in this section.

Define $\boldsymbol{\lambda}=[\lambda_1,\cdots,\lambda_m]$, and assume that $\lambda_1<\cdots<\lambda_m$ and that $[\lambda_1,\cdots,\lambda_m]$ are co-prime. As in [27], when the variance of the measurement noise is proportional to the wavelength, the probability of correct reconstruction of the integer vector $\mathbf{N}$, $Pr(\hat{\mathbf{N}}=\mathbf{N})$, is largely impacted by the minimum wavelength in the set $[\lambda_1,\cdots,\lambda_m]$.

We define the circular shift function $S(\cdot,\cdot)$ of an array as

$$\begin{aligned} S(\boldsymbol{\lambda},0)&=[\lambda_1,\cdots,\lambda_m]\\ S(\boldsymbol{\lambda},1)&=[\lambda_m,\lambda_1,\cdots,\lambda_{m-1}] \end{aligned}$$

$$\vdots$$

$$S(\boldsymbol{\lambda}, m-1) = [\lambda_2, \lambda_3, \lambda_4, \cdots, \lambda_m, \lambda_1]$$

Accordingly, recalling the definition of $\mathbf{A}(\cdot)$ and $\mathbf{B}(\cdot)$ in (4.64) and (4.65), we let $\mathbf{A}_j \triangleq \mathbf{A}(S(\boldsymbol{\lambda}, j)), \mathbf{B}_j \triangleq \mathbf{B}(S(\boldsymbol{\lambda}, j)), j = 0, 1, \cdots, m-1$.

As shown in Lemma 4.4, in the conventional lattice algorithm, the use of different basis $\mathbf{A}_j, j = 0, \cdots, m-1$ implies different reconstruction probability $Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, j)\right)$.

Lemma 4.4. Define $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]$. Suppose $\lambda_1 < \cdots < \lambda_m$ and are co-prime. Then the probabilities of correct reconstruction of the integer set $\mathbf{N}$ using the conventional lattice algorithm[27] satisfies

$$Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 0)\right) > Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, j)\right), j = 1, \cdots, m-1.$$

□

Proof. We only prove the case $Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 0)\right) > Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 1)\right)$. The proofs of other cases are similar.

As indicated in [27], for $S(\boldsymbol{\lambda}, 0)$, the lattice algorithm will return the true value if $|\omega_i - \omega_m| < \frac{1}{2}$ for $i = 1, \cdots, m-1$, and this yields

$$Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 0)\right) = Pr\left(\bigcap_{i=1}^{m-1} |\omega_i - \omega_m| < 1/2\right)$$

Similarly, for $S(\boldsymbol{\lambda}, 1)$, we have

$$\begin{aligned} & Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 0)\right) \\ = & Pr\Big(\bigcap_{i \in \{1, \cdots, m\} \setminus \{m-1\}} |\omega_i - \omega_{m-1}| < 1/2\Big) \end{aligned} \tag{4.67}$$

Since $\omega_i \sim \mathcal{N}(0, \delta^2 \lambda_i^2), i = 1, \cdots, m$, indicating that the variance of ω_i is proportional to the value of wavelengths λ_i, it is not hard to see that $Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 0)\right) > Pr\left(\hat{\mathbf{N}} = \mathbf{N} | S(\boldsymbol{\lambda}, 1)\right)$ □

The result in Lemma 4.4 is from a statistical perspective. However, in each realization, it is possible that the conventional lattice algorithm may return the wrong estimate of $\mathbf{N}$ using $\mathbf{A}_0$ while the algorithm returns the correct estimate using $\mathbf{A}_j, j = 1, \cdots, m-1$. Thus it is useful to draw information not only from $\mathbf{A}_0$, as used in conventional lattice algorithm in [27], but also from $\mathbf{A}_j, j = 1, \cdots, m-1$ before a decision is made. Based on the above idea, we present following estimator. The performance of the estimator is analysed in the next section.

$$\begin{aligned} \hat{\mathbf{N}} =& \hat{\mathbf{N}}_{j_0} \\ \text{s.t. } j_0 =& \arg \min_{j=0,\cdots,m-1} \left| \mathbb{E}\left[(\hat{\hat{r}}_j - \hat{\mathbf{r}}_j) \cdot \bar{\boldsymbol{\lambda}}_j \right] \right| \\ \hat{\mathbf{N}}_j =& \arg \min_{\mathbf{n} \in \mathbb{Z}^m} \|\mathbf{n}\mathbf{B}_j + \bar{\mathbf{y}}_i \mathbf{A}_j\|, j = 0, \cdots, m-1 \\ \hat{\mathbf{r}}_j =& \hat{\mathbf{N}}_j \boldsymbol{\lambda} + \mathbf{y}_j \\ \hat{\hat{r}}_j =& \mathbf{E}\left[r | \hat{\mathbf{r}}_j \right] \end{aligned} \tag{4.68}$$

where $\mathbb{E}[\cdot]$ is the arithmetic mean operator, $\bar{\boldsymbol{\lambda}}_j \triangleq S(\bar{\boldsymbol{\lambda}}, j)$, $\mathbf{y}_j \triangleq S(\mathbf{y}, j)$, $\bar{\mathbf{y}}_j = S(\bar{\mathbf{y}}, j)$, and $\hat{\hat{r}}_j = \mathbf{E}\left[r | \hat{\mathbf{r}}_j \right]$ is the optimal estimator of r given $\hat{\mathbf{r}}_j$, as can be found in [27].

It is not hard to solve (4.68). One could first solve $\hat{\mathbf{N}}_j, j = 0, \cdots, m-1$ according to $\mathbf{A}_j = \mathbf{A}(S(\boldsymbol{\lambda}, j))$ using the conventional lattice algorithm, and then compute the $\hat{\mathbf{r}}_j$. After calculating $\hat{\mathbf{N}}_j$ and $\hat{\mathbf{r}}_j$ for all j, the estimation of $\hat{\mathbf{N}}$ is easy to find by computing $\left| \mathbb{E}\left[(\hat{\hat{r}}_j - \hat{\mathbf{r}}_j) \cdot \bar{\boldsymbol{\lambda}}_j \right] \right|$.

The algorithm can be divided into two parts: the off-line part and the on-line part. The matrices used are constructed and stored in memory in the off-line part. In the on-line part, the unknown range is estimated using measurements and the stored matrices. The algorithm is set out as Algorithm 4.2.

4.4.3 The Computational Complexity and Performance Analysis

In this section, we aim to analyse the algorithm described in Algorithm 1 in terms of computational complexity and reconstruction performance. The computational complexity is measured by the required on-line computation time, while the reconstruction perfor-

Algorithm 4.2: Proposed algorithm to find $\hat{\mathbf{N}}$ with different constructions of $\mathbf{A}_j$

Data: $\boldsymbol{\lambda} = [\lambda_1, \cdots, \lambda_m]$,
$\lambda_1 < \cdots < \lambda_m$, $\{\lambda_1, \cdots, \lambda_m\}$ are co-prime,
$\mathbf{y} = [y_1, \cdots, y_m]$
Result: Estimation of $\mathbf{N}$, i.e. $\hat{\mathbf{N}}$

Off-line part:
for $j = 0 : m-1$ **do**
    $\mathbf{A}_j = \mathbf{A}(S(\boldsymbol{\lambda}, j))$;
    Compute $\mathbf{B}_j$ and other associated matrices;
On-line part:
$\bar{\mathbf{y}} = [y_1/\lambda_1, \cdots, y_m/\lambda_m]$;
for $j = 0 : m-1$ **do**
    $\mathbf{y}_j = S(\mathbf{y}, j)$, $\bar{\mathbf{y}}_j = S(\bar{\mathbf{y}}, j)$;
    Calculate $\hat{\mathbf{N}}_j$ using the conventional lattice algorithm with data $\bar{\mathbf{y}}_j\mathbf{A}_j$, $\mathbf{B}_j$ and other associated matrix;
    $\hat{\hat{r}}_j = E\left[r|\hat{\mathbf{r}}_j\right]$, $\text{Obj}_j = \left|\mathbb{E}\left[(\hat{\hat{r}}_j - \hat{\mathbf{r}}_j)\cdot\bar{\boldsymbol{\lambda}}_j\right]\right|$;
$j_0 = \arg\min_{j=0,\cdots,m-1} \text{Obj}_j$;
return $\hat{\mathbf{N}} = \hat{\mathbf{N}}_{j_0}$

mance is measured by the probability of correct reconstruction of $\mathbf{N}$; that is, $Pr(\hat{\mathbf{N}} = \mathbf{N})$.

The required on-line computation has two parts, 1) the processes to calculate the candidate estimations, i.e. $\hat{\mathbf{N}}_j, j = 0, \cdots, m-1$, and 2) the outer loop from 0 to $m-1$. Since, for each j, the first part needs m times computations and therefore, the total computation complexity is $m \cdot m = m^2$.

Since the performance of the proposed algorithm is measured by $Pr(\hat{\mathbf{N}} = \mathbf{N})$, as indicated in the algorithm, $\hat{\mathbf{N}} = \mathbf{N}$ happens if and only if there exists at least one $j_0 \in [0, \cdots, m-1]$ such that following event happens

$$\{\hat{\mathbf{N}}_{j_0} = \mathbf{N}\} \bigcap \{j_0 = \arg \min_{j=0,\cdots,m-1} \text{Obj}_j\}, \tag{4.69}$$

where $\text{Obj}_j = \left|\mathbb{E}\left[(\hat{\hat{r}}_j - \hat{\mathbf{r}}_j)\cdot\bar{\boldsymbol{\lambda}}_j\right]\right|$. Therefore

$$\begin{aligned} & Pr\left(\hat{\mathbf{N}} = \mathbf{N}\right) \\ =& Pr\left(\{\hat{\mathbf{N}}_{j_0} = \mathbf{N}\} \bigcap \{j_0 = \arg \min_{j=0,\cdots,m-1} \text{Obj}_j\}\right) \end{aligned}$$

$$=Pr\left(j_0=\arg\min_{j=0,\cdots,m-1}\mathrm{Obj}_j\middle|\hat{\mathbf{N}}_{j_0}=\mathbf{N}\right)Pr\left(\hat{\mathbf{N}}_{j_0}=\mathbf{N}\right)$$

Since, when $\hat{\mathbf{N}}_{j_0}=\mathbf{N}$, we have $\hat{\hat{r}}_{j_0}\simeq r$,

$$(\hat{\hat{r}}_{j_0}-\hat{\mathbf{r}}_{j_0})\cdot\bar{\boldsymbol{\lambda}}_{j_0}\simeq(r-\hat{\mathbf{r}}_{j_0})\cdot\bar{\boldsymbol{\lambda}}_{j_0}=\boldsymbol{\omega}_{j_0}\cdot\bar{\boldsymbol{\lambda}}_{j_0},$$

where $\boldsymbol{\omega}_{j_0}\triangleq S(\boldsymbol{\omega},j_0)$.

According to the definition of $\boldsymbol{\omega}_j$ and $\bar{\boldsymbol{\lambda}}_{j_0}$, we know that

$$\boldsymbol{\omega}_{j_0}\cdot\bar{\boldsymbol{\lambda}}_{j_0}\sim\mathcal{N}(0,\delta^2\mathbf{I}_m)$$

where $\mathbf{I}_m$ is the m-dimensional identity matrix.

If $\hat{\mathbf{N}}_{j_0}=\mathbf{N}$, then $\left|\mathbb{E}\left[(\hat{\hat{r}}_{j_0}-\hat{\mathbf{r}}_{j_0})\cdot\bar{\boldsymbol{\lambda}}_{j_0}\right]\right|\simeq 0$, and then $\left|\mathbb{E}\left[(\hat{\hat{r}}_{j_0}-\hat{\mathbf{r}}_{j_0})\cdot\bar{\boldsymbol{\lambda}}_{j_0}\right]\right|$ will normally be the minimum value for $j=0,\cdots,m-1$ if m is large.

This implies that

$$\begin{aligned}&Pr\left(j_0=\arg\min_{j=0,\cdots,m-1}\mathrm{Obj}_j\middle|\hat{\mathbf{N}}_{j_0}=\mathbf{N}\right)\\ \simeq&Pr\left(j_0=\arg\min_{j=0,\cdots,m-1}\mathrm{Obj}_j\middle|\hat{\mathbf{N}}_{j_0}=\mathbf{N},\mathrm{Obj}_{j_0}=0\right)\\ =&1\end{aligned}$$

It follows that $Pr\left(\hat{\mathbf{N}}=\mathbf{N}\right)\simeq Pr\left(\hat{\mathbf{N}}_{j_0}=\mathbf{N}\right)$. Define the event $\mathcal{E}\triangleq\{\hat{\mathbf{N}}_{j_0}=\mathbf{N},\exists j_0\in\{0,\cdots,m-1\}\}$, then $\mathcal{E}^c$ states that $\{\hat{\mathbf{N}}_j\neq\mathbf{N},\forall j=0,\cdots,m-1\}$. In consequence,

$$\begin{aligned}Pr(\mathcal{E})=1-Pr(\mathcal{E}^c)=&1-Pr\left(\bigcap_{j=0}^{m-1}\{\hat{\mathbf{N}}_j\neq\mathbf{N}\}\right)\\ >&1-Pr\left(\hat{\mathbf{N}}_j\neq\mathbf{N}\right)\\ =&Pr\left(\hat{\mathbf{N}}_j=\mathbf{N}\right)\quad\forall j=0,\cdots,m-1\end{aligned}$$

By Lemma 4.4, this implies that the proposed algorithm is more robust than the conventional lattice algorithm in terms of the probability of correct reconstruction.

4.4.4 Simulation

The performance using the proposed algorithm listed in Algorithm 4.2 is demonstrated using Monte Carlo simulations. We take the US UHF band, 902 $\sim$ 928 MHz, as an example. The frequency 902 MHz is selected as the reference frequency; that is, $f_0 = 902$ MHz, and therefore, the available frequency differences are within 1 $\sim$ 26 MHz with 1MHz increment[82] and wavelengths are within 300 $\sim$ 12 m. In the simulation, the following co-prime wavelengths are selected $\boldsymbol{\lambda} = \{21,\ 19,\ 17,\ 13\}$m.

The distance r is randomly selected between 1 and the LCM of all . The parameter δ in the measurement noise variance $\delta^2\lambda_i^2$ is chosen so that $-20\log_{10}\delta = 34 : 2 : 50$, which provides an indication for both noise level and signal-to-noise ratio in the simulation. All simulation results illustrated are averaged over 5000 Monte Carlo runs. The algorithm performance is evaluated in terms of the reconstruction probability $Pr(\mathbf{N} = \hat{\mathbf{N}})$ versus the phase measurement noise level.

The simulation results are shown in Figure 4.7. For a comparison, the performance of the conventional lattice algorithm using the lattice basis $\mathbf{A}_0$ is presented as well. As indicated in Lemma 4.4, the reconstruction probability of the conventional lattice algorithm satisfies $Pr(\hat{\mathbf{N}} = \mathbf{N}|S(\boldsymbol{\lambda}, 0)) > Pr(\hat{\mathbf{N}} = \mathbf{N}|S(\boldsymbol{\lambda}, j)), j = 1, \cdots, m-1$ and $S(\boldsymbol{\lambda}, 0)$ corresponds to the lattice basis $\mathbf{A}_0$. Thus it is sufficient to compare the performance of the new algorithm and conventional lattice algorithm using the lattice basis $\mathbf{A}_0$.

Figure 4.7 shows that the proposed algorithm outperforms the conventional lattice algorithm in that it has a high probability of correct reconstruction for a given noise level.

4.4.5 Conclusions

In this section, we have presented an improved lattice algorithm for estimating the distance. The proposed algorithm is more robust than the conventional lattice algorithm in terms of the probability of correct reconstruction of the wrapped integers at the cost of only slightly higher complexity. The performance of the proposed algorithm is analysed. This algorithm is suitable for other situations when wrapped measurements arise.

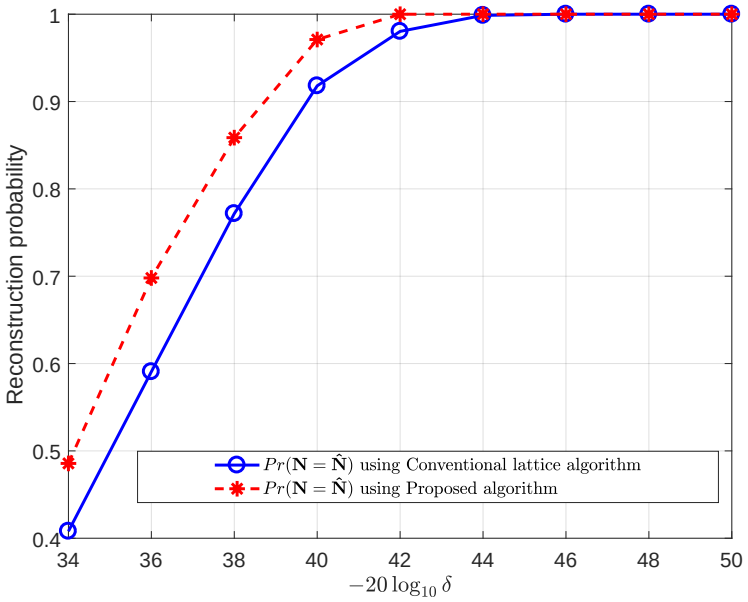


Figure 4.7: The reconstruction probability comparison between the conventional lattice [10] algorithm and proposed algorithm in Section 4.4.2.

4.5 An Improved Lattice Algorithm Using Multiple Candidate Lattice Points

4.5.1 Introduction

As mentioned earlier, in many engineering applications, e.g. sensor localization using range, the performance of successfully reconstructing the ambiguous integers is the first priority since if the integers are not correctly estimated, the range estimation will be normally far away from the ground truth. In this section, we present an improved lattice algorithm by considering multiple lattice points with slightly more computations comparing to the conventional lattice algorithm in [27]. In the conventional algorithm, due to the noise, the given point corresponding to the measurements may lie in the Voronoi cell of a neighbor of the ground truth and, in this situation, an incorrect estimate will be returned. Base on this, a set of neighbor lattice points of the one given by conventional algorithm is considered, and then a point is selected within these points as the estimation via a cost function. The improved algorithm is demonstrated to have higher probability to reconstruct the ambiguous integers. The performance of the improved algorithm is

analysed as well.

4.5.2 The Proposed Algorithm

Suppose that the wavelengths used are co-prime, then the estimator can be written as:

$$\hat{\mathbf{N}} = \arg\min_{\mathbf{n}\in\mathbb{Z}^m} \left\| \mathbf{nB} + \bar{\mathbf{y}}\mathbf{A} \right\|. \tag{4.70}$$

where

$$\mathbf{A} = \begin{pmatrix} \prod_{i=1}^{m-1}\lambda_i & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \prod_{i=1}^{m-1}\lambda_i & 0 \\ -\frac{\prod_{i=1}^{m}\lambda_i}{\lambda_1} & \dots & -\frac{\prod_{i=1}^{m}\lambda_i}{\lambda_{m-1}} & 0 \end{pmatrix} \tag{4.71}$$

and,

$$\mathbf{B} = \begin{pmatrix} \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_1} & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & \frac{\prod_{i=1}^{m-1}\lambda_i}{\lambda_{m-1}} & 0 \\ 0 & \dots & 0 & 0 \end{pmatrix} \tag{4.72}$$

In formula (4.70), $\mathbf{nB}$ generates a lattice space, denote by $\mathbf{\Omega}(\mathbf{B})$ with basis $\mathbf{B}$. On the other hand, each lattice point, $\mathbf{nB}$, in $\mathbf{\Omega}(\mathbf{B})$ has a Voronoi cell, denoted by $\boldsymbol{V}(\mathbf{nB})$, and these Voronoi cells are translated of each other.

In formula (4.70), we find that $\bar{\mathbf{y}}$ can be written

$$\begin{aligned} \bar{\mathbf{y}} &= \left[\frac{\omega_1}{\lambda_1}, \cdots, \frac{\omega_m}{\lambda_m}\right] + \left[\frac{y_{1,true}}{\lambda_1}, \cdots, \frac{y_{m,true}}{\lambda_m}\right] \\ &\triangleq \bar{\boldsymbol{\omega}} + \bar{\mathbf{y}}_{true} \end{aligned}$$

where $\bar{\boldsymbol{\omega}} \sim \mathcal{N}(\mathbf{0}, \delta^2\mathbf{I})$. Then

$$\bar{\mathbf{y}}\mathbf{A} = \bar{\boldsymbol{\omega}}\mathbf{A} + \bar{\mathbf{y}}_{true}\mathbf{A}. \tag{4.73}$$

(4.73) implies that the given point in (4.70) can be written as a sum of the true value and an error. It is easy to see that $\bar{\omega}\mathbf{A}$ is still Normally distributed, i.e. $\bar{\omega}\mathbf{A} \sim \mathcal{N}(0, \mathbf{\Sigma})$, where

$$\mathbf{\Sigma} = \delta^2 \mathbf{A}_1^T \mathbf{A}_1 \tag{4.74}$$

and $\mathbf{A}_1$ is just the first $m-1$ columns of $\mathbf{A}$. Accordingly,

$$\bar{\mathbf{y}}\mathbf{A} \sim \mathcal{N}(\bar{\mathbf{y}}_{true}\mathbf{A}, \mathbf{\Sigma}) \tag{4.75}$$

The conventional lattice algorithm gives the closest point to the $\bar{\mathbf{y}}\mathbf{A}$ in the lattice space generated by $\mathbf{B}$ and this closest point corresponds to the estimation of the $\hat{\mathbf{N}}$. If there is no noise in $\bar{\mathbf{y}}$, i.e. $\bar{\mathbf{y}}\mathbf{A} = \bar{\mathbf{y}}_{true}\mathbf{A}$, the given point will be exactly a point in the lattice. However, as shown in (4.75), $\bar{\mathbf{y}}\mathbf{A}$ is Normally distributed and this defines a ellipsoid in the Euclidean space centred at the true point $\bar{\mathbf{y}}_{true}\mathbf{A}$ with area determined by the covariance $\mathbf{\Sigma}$ and the confidence level α.

Given a point $\bar{\mathbf{y}}\mathbf{A}$, we known that this point can be in any Voronoi cell covered by the ellipsoid. Intuitively, the conventional lattice algorithm will not return the true lattice point if the $\bar{\mathbf{y}}\mathbf{A}$ lies in other Voronoi cells. Thus, if the given point $\bar{\mathbf{y}}\mathbf{A}$ does not lie in the $\boldsymbol{V}(\bar{\mathbf{y}}_{true}\mathbf{A})$, then the lattice algorithm will return a closest point $\hat{\mathbf{N}}\mathbf{B}$ which is not equal to $\bar{\mathbf{y}}_{true}\mathbf{A}$. Since $\mathbf{B}$ is a diagonal matrix and $\hat{\mathbf{N}}\mathbf{B} \in \boldsymbol{\Omega}(\mathbf{B})$, then there exists a integer vector $\mathbf{t}_0 = [t_{0,1}, \cdots, t_{0,m-1}, 0] \in \mathbb{Z}^{m-1}$ such that

$$\mathbf{t}_0\mathbf{B} + \hat{\mathbf{N}}\mathbf{B} = \bar{\mathbf{y}}_{true}\mathbf{A} \tag{4.76}$$

Obviously, if $\hat{\mathbf{N}}\mathbf{B} = \bar{\mathbf{y}}_{true}\mathbf{A}$, $\mathbf{t}_0 = 0$. We aim to estimate the unknown vector $\mathbf{t}_0 \in \mathbb{Z}^{m-1}$ in order to improve the performance of the lattice algorithm.

From the above analysis and the distribution of $\bar{\mathbf{y}}\mathbf{A}$, we know that $\bar{\mathbf{y}}_{true}\mathbf{A}$ either equals $\hat{\mathbf{N}}\mathbf{B}$ or is a lattice point which is a neighbor of $\hat{\mathbf{N}}\mathbf{B}$ within a certain distance determined by δ^2. In practice, δ^2 is often small, thus we assume that $|t_{0,i}| \leq 1$ and $t_{0,i} \in \mathbb{Z}$ for $i = 1, \cdots, m-1$. Let $\mathbf{t}_j$, $j = 1, \cdots, j_{max}$, be the possible value of $\mathbf{t}_0$, i.e.

$$\mathbf{t}_j = [t_{j,1}, \cdots, t_{j,m-1}, 0], \text{ s.t. } |t_{j,i}| \leq 1 \text{ and } t_{j,i} \in \mathbb{Z} \tag{4.77}$$

and it is easily to determine that $j_{max} = 3^{m-1}$. From (4.76), we could define a collection of all possible neighboring points of $\hat{\mathbf{N}}\mathbf{B}$ and denoted by $\mathcal{P}$, i.e.

$$\mathcal{P} = \{\hat{\mathbf{N}}\mathbf{B} + \mathbf{t}_1\ \mathbf{B}, \cdots, \hat{\mathbf{N}}\mathbf{B} + \mathbf{t}_{j_{max}}\ \mathbf{B}\}$$

The number of points included in $\mathcal{P}$ is denoted by $\#\mathcal{P}$. From the above analysis, we can see that $\#\mathcal{P}$ contains $\mathbf{t}_0$ with large probability. It should be noticed that a larger region for $|t_{0,i}|$ can be used, however, it needs more computations in estimating $\mathbf{t}_0$. Obviously, the zero vector $\mathbf{0}$ is a possible vector of $\mathbf{t}_j$, therefore, without loss of generality, we assume that $\mathbf{t}_{\lfloor \frac{j_{max}}{2} \rceil} = \mathbf{0}$.

$\mathcal{P}$ can be used directly to estimate $\mathbf{t}_0$, however, we can reduce the number of elements in $\mathcal{P}$ by computing the lattice points covered by the ellipsoid via the Mahalanobis distance:

$$\begin{aligned} d_j &= (\bar{\mathbf{y}}_0\mathbf{A} + \mathbf{t}_j - \bar{\mathbf{y}}_0\mathbf{A})\delta^2\mathbf{\Sigma}(\bar{\mathbf{y}}_0\mathbf{A} + \mathbf{t}_j - \bar{\mathbf{y}}_0\mathbf{A})^T \\ \Rightarrow d_j &= \delta^2(\mathbf{t}_j\mathbf{B})\mathbf{\Sigma}(\mathbf{t}_j\mathbf{B})^T \end{aligned} \tag{4.78}$$

Then, in the (4.78), the index of the points within the ellipsoid can be calculated by

$$\mathcal{I} = \{j; d_j < \chi^2_\alpha\} \tag{4.79}$$

suppose $\mathcal{P}_1$ contains the lattice points in the $\mathcal{P}$ satisfying (4.79) and the number of points included in $\mathcal{P}_1$ is denoted by $\#\mathcal{P}_1$. It should be noticed that $\mathcal{I}$ can be calculated in advance.

An 2-D example of how to construct $\mathcal{P}$ is shown in Fig. 4.8, where $\mathbf{R}$ is the region including all possible point in $\mathcal{P}$ and $\mathbf{E}$ is the ellipsoid defined by $\mathcal{N}(0, \mathbf{\Sigma})$. In Fig.4.8, the given point lies in the Voronoi cell which is not of the true point, then the lattice algorithm will give the wrong estimation, i.e. $\hat{\mathbf{N}}\mathbf{B}$. If we can estimate $\mathbf{t}_0\mathbf{B} = \hat{\mathbf{N}}\mathbf{B} - \mathbf{n}\mathbf{B}$ correctly, then the true estimate is achieved.

In Fig.4.8, the given point lies in the Voronoi cell which is not that of the true point, then the lattice algorithm will give the wrong estimation, i.e. $\hat{\mathbf{N}}\mathbf{B}$. If we can estimate

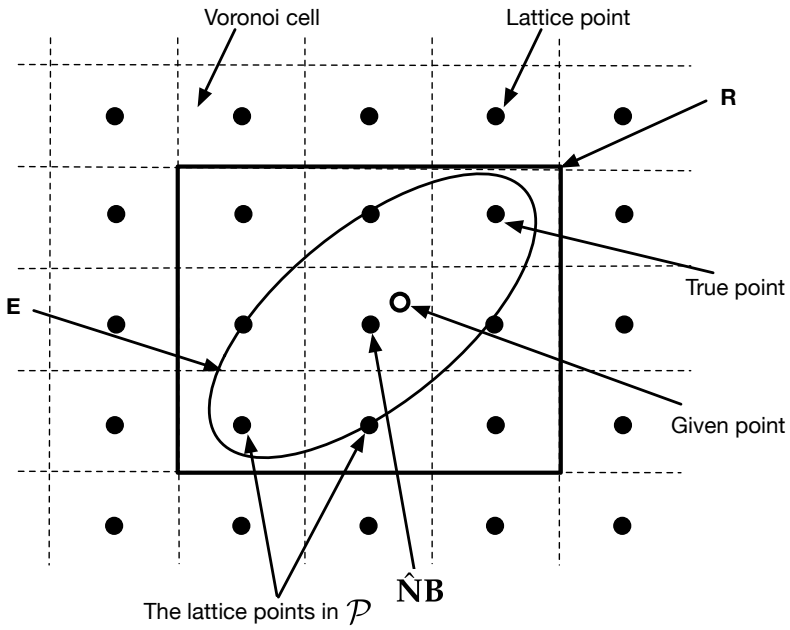


Figure 4.8: The example of $\mathcal{P}$

$\mathbf{t}_0\mathbf{B} = \hat{\mathbf{N}}\mathbf{B} - \mathbf{n}\mathbf{B}$ correctly, then the true estimation is achieved.

Now the main goal is to find the estimation of $\mathbf{t}_0$. An efficient algorithm is presented as follows:

$$\hat{j}_0 = \arg\min_{j\in\mathcal{I}} \frac{1}{m}\sum_{i=1}^{m}\hat{\omega}_{j,i}^2 \tag{4.80}$$

$$s.t. \quad \hat{\mathbf{N}}_{\left\lfloor \frac{j_{max}}{2} \right\rceil} = \textbf{Lattice Algorithm}(\boldsymbol{\lambda}, \mathbf{y}) \tag{4.81}$$

$$\hat{\mathbf{N}}_j = \hat{\mathbf{N}}_{\left\lfloor \frac{j_{max}}{2} \right\rceil} + \mathbf{t}_j, \forall j \in \{1, \cdots, j_{max}\} \setminus \left\{ \left\lfloor \frac{j_{max}}{2} \right\rceil \right\} \tag{4.82}$$

$$\hat{\tilde{\omega}}_{j,i} = \frac{\lfloor \hat{r}_j \rceil - (\hat{N}_{j,i}\lambda_i + y_i)}{\lambda_i}, \forall i = 1, \cdots, m \tag{4.83}$$

$$\hat{r}_j = \mathbf{E}[r|\hat{\mathbf{N}}_j] \tag{4.84}$$

where $\textbf{Lattice Algorithm}(\boldsymbol{\lambda}, \mathbf{y})$ is the conventional lattice algorithm using $\boldsymbol{\lambda}$ and $\mathbf{y}$ and $\mathbf{E}[r|\hat{\mathbf{N}}_j]$ is the estimate of r, which can be done using a maximum likelihood method [27], i.e

$$\hat{r} = W\sum_{i=1}^{k}(\hat{N}_i\lambda_i + y_i)W_i \tag{4.85}$$

where $W_i = 1/\lambda_i^2$, $W = 1/\sum_{i=1}^{k} W_i$.

4.5.3 Performance Analysis

Next, we aim to analyze the performance of the algorithm we have presented under some mild assumptions. Suppose that there exists an index $j_0 \in \{1, \cdots, 3^{m-1}\}$ satisfying $\hat{N}_{j_0,m} = n_m$. Let $\mathcal{J} = \{1, \cdots, 3^{m-1}\} \setminus j_0$, then Algorithm (4.80) returns the true value, i.e. $\hat{j}_0 = j_0$, if

$$\frac{1}{m}\sum_{i=1}^{m} \hat{\omega}_{j_0,i}^2 - \frac{1}{m}\sum_{i=1}^{m} \hat{\omega}_{j,i}^2 < 0, \quad \forall j \in \mathcal{J} \tag{4.86}$$

Consider $\lambda_i \hat{\omega}_{j_0,i}$, we have

$$\begin{aligned}
\lambda_i \hat{\omega}_{j_0,i} &= \lfloor \hat{r}_{j_0} \rceil - \hat{N}_{j_0,i}\lambda_i - y_{i,true} - \omega_i \\
&= \left\lfloor r + W\textstyle\sum_{i=1}^{k} W_i \omega_i \right\rceil - r - \omega_i \\
&\triangleq r + I_{j_0,i} - r - \omega_i, \quad \text{where} \quad I_{j_0,i} \in \mathbb{Z} \\
&= I_{j_0,i} - \omega_i
\end{aligned} \tag{4.87}$$

since δ^2 is small and ω_i is normally distributed with 0 mean, then $I_{j_0,i} = 0$ with high probability, therefore, we assume that $I_{j_0,i} = 0$ in the following analysis.

For a fixed $j \in \mathcal{J}$, since there exits $\alpha_{j,i} \in \mathbb{Z}, i = 1, \cdots, m$ such that $\hat{N}_{j,i}\lambda_i + y_{i,true} + \alpha_{j,i}\lambda_i = r$ and at least one $\alpha_{j,i} \neq 0$, otherwise, $j = j_0$. Therefore,

$$\begin{aligned}
\lambda_i \hat{\omega}_{j,i} &= \left\lfloor W\textstyle\sum_{i=1}^{m}(r - \alpha_{j,i}\lambda_i)W_i + \bar{\omega} \right\rceil - r - \alpha_i\lambda_i - \omega_i \\
&= \left\lfloor r - W\textstyle\sum_{i=1}^{m} \alpha_{j,i}\lambda_i W_i + \bar{\omega} \right\rceil - r - \alpha_i\lambda_i - \omega_i \\
&\triangleq \lfloor r - \bar{\alpha}_j + \bar{\omega} \rceil - r - \alpha_i\lambda_i - \omega_i
\end{aligned} \tag{4.88}$$

Proposition 4.6. For a fixed $j \in \mathcal{J}$, there exists at least an $i \in \{1, \cdots, m\}$ such that $\lfloor r - \bar{\alpha}_j + \bar{\omega} \rceil - r - \alpha_i\lambda_i \neq 0$. □

Proof. Since the $\bar{\alpha}_j = W\sum_{i=1}^{m}\alpha_{j,i}\lambda_i W_i \in \mathbb{R}$, $\bar{\alpha}_{j,i} \in \mathbb{R}$ and can be written as a sum of the integer part $\bar{\alpha}_j^I \in \mathbb{Z}$ and fractional part $\bar{\alpha}_j^D \in [-1, 1)$. Therefore,

$$\begin{aligned}
&\left\lfloor r - \bar{\alpha}_j + \bar{\omega} \right\rceil - r - \alpha_i\lambda_i \\
=&r - \bar{\alpha}_j^I + \left\lfloor \bar{\omega} - \bar{\alpha}_j^D \right\rceil - r - \alpha_i\lambda_i \\
=&\left\lfloor \bar{\omega} - \bar{\alpha}_j^D \right\rceil - \bar{\alpha}_j^I - \alpha_i\lambda_i
\end{aligned} \tag{4.89}$$

where $\left\lfloor \bar{\omega} - \bar{\alpha}_j^D \right\rceil - \bar{\alpha}_j^I$ is a constant for a fixed j. If $\left\lfloor \bar{\omega} - \bar{\alpha}_j^D \right\rceil - \bar{\alpha}_j^I - \alpha_i\lambda_i = 0$ for all $i = 1, \cdots, m$, then we must have $\alpha_1\lambda_1 = \cdots = \alpha_m\lambda_m$ which implies that $\alpha_1 = \cdots = \alpha_m = 0$ as $\alpha_m \in \mathbb{Z}$ and λ_i are not equal to each other. This contradicts the fact that there exists at least one $\alpha_{j,i} \neq 0$. □

Let $I_{j,i} \triangleq \left\lfloor r - \bar{\alpha}_j + \bar{\omega} \right\rceil - r - \alpha_i\lambda_i$. From Prop. 4.6, we know $I_{j,i} \in \mathbb{Z}$ and $\sum_{i=1}^{k} I_{j,i}^2 \in \mathbb{Z}^+$. Then from (4.88), we have $\hat{\tilde{\omega}}_{j,i} = \frac{I_{j,i}}{\lambda_i} - \frac{\omega_i}{\lambda_i}$. This with (4.86) adnd (4.87) implies that

$$\begin{aligned}
&\frac{1}{m}\left(\sum_{i=1}^{m}\hat{\tilde{\omega}}_{j_0,i}^2 - \sum_{i=1}^{m}\hat{\tilde{\omega}}_{j,i}^2\right) \\
=&\frac{1}{m}\sum_{i=1}^{m}\left(-\frac{\omega_i^2}{\lambda_i^2} - \left(\frac{I_{j,i}}{\lambda_i} - \frac{\omega_i}{\lambda_i}\right)^2\right) \\
\triangleq&\frac{1}{m}\left(-\sum_{i=1}^{m}\frac{I_{j,i}^2}{\lambda_i^2} + 2\sum_{i=1}^{m}\tilde{\omega}_i\right), \quad \tilde{\omega}_i = \frac{\omega_i}{\lambda_i} \sim \mathcal{N}(0, \delta^2)
\end{aligned}$$

In consequence, the probability that the proposed algorithm returns the true value is

$$Pr\left(\bigcap_{j\in\mathcal{J}}\left\{-\sum_{i=1}^{m}\frac{I_{j,i}^2}{\lambda_i^2} + 2\sum_{i=1}^{m}\frac{\tilde{\omega}_i}{\lambda_i} < 0\right\}\right) \tag{4.90}$$

The calculation of (4.90) is trivial, but it is possible to find an approximation. Since in range estimation, the $\lambda_i, i = 1, \cdots, m$ are from the same frequency band and close to each other. We could use $\tilde{\lambda} = \frac{1}{n}\sum_{i=1}^{m}\lambda_i$ to approximate λ_i. Therefore, (4.90) becomes

$$Pr\left(\bigcap_{j\in\mathcal{J}}\left\{-\frac{1}{\tilde{\lambda}}\sum_{i=1}^{m}I_{j,i}^2 + \sum_{i=1}^{k}2I_{j,i}\tilde{\omega}_i < 0\right\}\right) \tag{4.91}$$

and

$$\left(-\frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} I_{j,i}^2 + \sum_{i=1}^{k} 2I_{j,i}\tilde{\omega}_i\right) \sim \mathcal{N}\left(-\frac{\sum_{i=1}^{m} I_{j,i}^2}{\tilde{\lambda}}, 4\delta^2 \sum_{i=1}^{m} I_{j,i}^2\right)$$

when δ is small, $Pr\left(-\frac{1}{\tilde{\lambda}}\sum_{i=1}^{m} \tilde{I}_{j,i}^2 + \sum_{i=1}^{k} 2I_{j,i}\tilde{\omega}_i < 0\right)$ is largely influenced by $\sum_{i=1}^{m} I_{j,i}^2 = 1$ because this probability is closed to 1 for other values of $\sum_{i=1}^{m} I_{j,i}^2$. Therefore, we could approximate (4.91) by only considering the case $\sum_{i=1}^{m} I_{j,i}^2 = 1$, i.e. only one element in $[I_{j,1}, \cdots, I_{j,m}]$ is ± 1 and other elements are all 0. Therefore, by considering all possible combinations of $[I_{j,1}, \cdots, I_{j,m}]$ satisfying $\sum_{i=1}^{m} I_{j,i}^2 = 1$, (4.91) can be calculated approximately by

$$\begin{aligned} &Pr\left(\bigcap_{i=1,\cdots,m}\left\{-\frac{1}{\tilde{\lambda}} + 2\tilde{\omega}_i < 0, -\frac{1}{\tilde{\lambda}} - 2\tilde{\omega}_i < 0\right\}\right) \\ =&\prod_{i=1}^{k} Pr\left(-\frac{1}{\tilde{\lambda}} + 2\tilde{\omega}_i < 0, -1/\tilde{\lambda} - 2\tilde{\omega}_i < 0\right) \end{aligned} \tag{4.92}$$

Define two events $\mathcal{E}_i^{\pm} = \left\{-\frac{1}{\tilde{\lambda}} \pm 2\tilde{\omega}_i\right\}$, then

$$\begin{aligned} &Pr\left(\mathcal{E}_i^{+} \cap \mathcal{E}_i^{-}\right) \\ =&Pr\left(\mathcal{E}_i^{+} | \omega_i < 0\right) Pr(\omega_i < 0) + Pr\left(\mathcal{E}_i^{-} | \omega_i > 0\right) Pr(\omega_i > 0) \\ =&Pr\left(-\frac{1}{\tilde{\lambda}} - 2|\tilde{\omega}_i| < 0\right) \end{aligned}$$

It follows that

$$\prod_{i=1}^{k} Pr\left(-\frac{1}{\tilde{\lambda}} + 2|\tilde{\omega}_i| < 0\right) \tag{4.93}$$

which is computable since $-\frac{1}{\tilde{\lambda}} + 2|\tilde{\omega}_i|$ satisfies a folded normal distribution. In the next section, the simulation demonstrates the efficiency of (4.93) in range estimation.

4.5.4 Simulation

The proposed algorithm in this section is evaluated via simulation using two sets of wavelengths:

$$\mathbf{\Lambda}_1 = \{133,\ 131,\ 127,\ 125,\ 123,\ 122,\ 121\}$$

$$\mathbf{\Lambda}_2 = \{61,\ 59,\ 51,\ 49,\ 47,\ 46\}$$

These two sets are from the Wifi 2.4GHz and 5GHz bands which are widely used in sensor localization.

The distance r is randomly selected between 1 and the LCM of all wavelengths used. The parameter δ in the measurement noise variance $\delta^2\lambda_i^2$ is chosen as $-20\log_{10}\delta = 46 : 2 : 66$, which provides an indication for both noise level and signal-to-noise ratio in the simulation. The simulation results are averaged over 5000 Monte Carlo runs. The algorithm performance is evaluated in terms of the reconstruction probability $Pr(N_k = \hat{N}_k|\delta^2, \mathbf{\Lambda}_{1,2})$ versus the phase measurement noise level.

In the figure, `algorithm2` represents the algorithm proposed in this section. The simulation results of the reconstruction probability via Monte Carlo simulation and theoretic computation using (4.93) are given in Fig.4.9 and Fig.4.10. It is found that the simulation results agree with the theoretic result. As comparison, the performance of existing algorithm, represented by `algorithm1`, is plotted along with that of the proposed algorithm in Fig.4.9. Clearly, when both waveform sets are used, the proposed algorithm outperforms `algorithm1` in terms of reconstruction probability of the underlying integer set when signal to noise ratio is low.

On the other hand, the comparison of the computation complexity using $\boldsymbol{\mathcal{P}}$ and $\boldsymbol{\mathcal{P}}_1$ in plotted in Fig.4.10 by the ratio of $\#\boldsymbol{\mathcal{P}}$ and $\#\boldsymbol{\mathcal{P}}_1$, i.e. $\frac{\#\boldsymbol{\mathcal{P}}_1}{\#\boldsymbol{\mathcal{P}}}$. Obviously, $\#\boldsymbol{\mathcal{P}}_1 \leq \#\boldsymbol{\mathcal{P}}$, then $\frac{\#\boldsymbol{\mathcal{P}}_1}{\#\boldsymbol{\mathcal{P}}} \leq 1$. The smaller the value of $\frac{\#\boldsymbol{\mathcal{P}}_1}{\#\boldsymbol{\mathcal{P}}}$, the less computation required.

4.5.5 Conclusions

In this section, an improved lattice is presented by considering multiple lattice points. The proposed algorithm has better performance than the conventional lattice algorithm

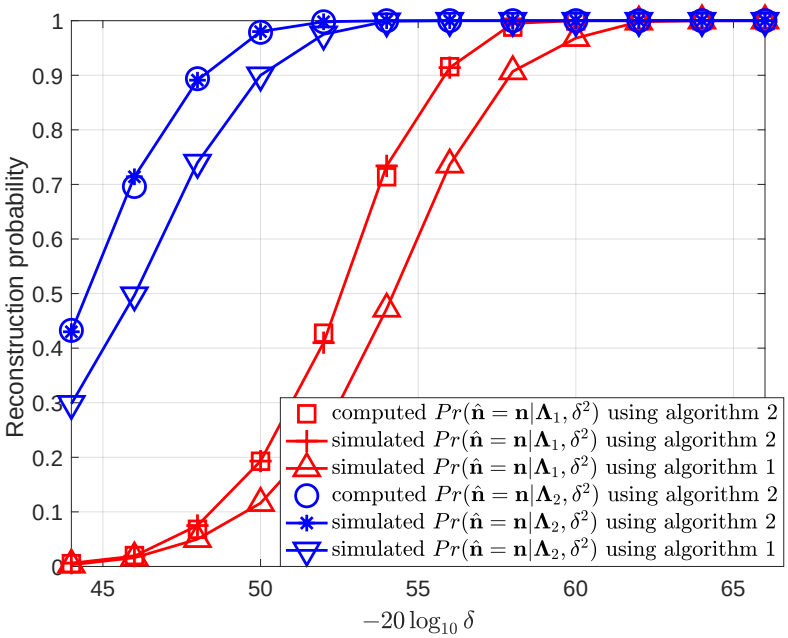


Figure 4.9: The comparison of reconstruction probability between proposed algorithm and conventional one

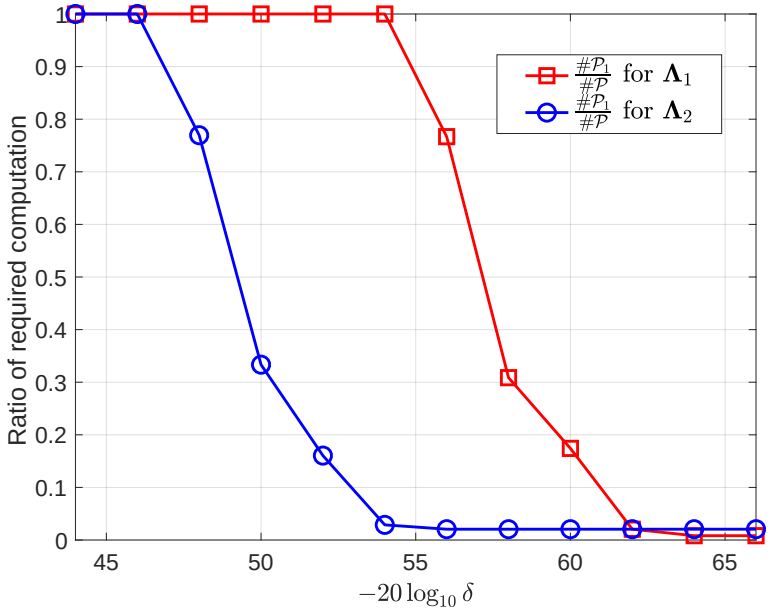


Figure 4.10: The comparison of the computation complexity

in terms of the probability of correct reconstruction with slightly higher complexity. The performance of the proposed algorithm is analysed and the approximated reconstruction probability is derived.

4.6 Conclusions

In this chapter, some improved algorithms based on lattice method are presented in order to fit more situations. Compared with conventional lattice method, the proposed relaxed condition can give more flexibility to designers on selecting wavelengths when the closed-form lattice algorithm is used while the other improved lattice algorithms focus on increasing reconstruction probability by using the different structures of basis, Voronoi cell and fast closest point searching algorithm etc.. The performance of the improved algorithms are evaluated and demonstrated via numerical simulations using various wavelengths setting. These improved algorithms can extend the applications of the lattice algorithm.

Part III

Finding A Position With Distance Dependent Noise

Chapter 5
Introduction and Preliminaries

In order to introduce the main results in Chapter 6 and 7, some useful theory, definitions and theorems, such as martingale, convergence and law of large number, are introduced in this chapter.

5.1 Introduction to Martingale Theory

Martingales are nowadays considered to be a powerful tool in solving problems in applied probability theory. Such problems include random walks, point processes, mathematical statistics, risk analysis and the mathematics of finance[83]. They also have applications in other fields of mathematics, such as harmonic analysis and partial differential equations[84, 85], economic [86] and engineering[87–89].

A martingale is the stochastic model of "fair game" in gambling which means a gambler has the same expectation to win or lose money. In [90], the theory of martingale is introduced in detail and the potential applications of martingale are revealed. Martingales are also widely used in proving the consistency of MLE as shown in [91–93]. Further introduction to martingale can also be found in [92] and [94].

In this section, basic facts about martingales are given and some useful definitions are introduced.

Let $(\mathbf{\Omega}, \mathcal{F}, (\mathcal{F}_i)_{i\geq 0}, P)$ be a filtered probability space: $\mathbf{\Omega}$ is a set, $\mathcal{F}$ is a σ-field of subset of Ω and P a probability measure defined on $\mathcal{F}$ and $\mathcal{F}_0 = \{\mathbf{\Omega}, \emptyset\}$.

Definition 5.1. A filtration is an increasing sequence of sub-σ-fields $\mathcal{F}_i, i > 0, i \in \mathbb{N}$ of $\mathcal{F}$, i.e. all $\mathcal{F}_i \subset \mathcal{F}$ are σ-fields and $\mathcal{F}_n \subset \mathcal{F}_m$ if $n \leq m$.

Given a process $\{X_i, i > 0, i \in \mathbb{N}\}$, X_i is a adapted to a filtration $\mathcal{F}_i$ if $X_i \in \mathcal{F}_i$, i.e. X_i is $\mathcal{F}_i$-measurable requiring $X_i^{-1}(\mathbb{B}) \in \mathcal{F}_i$ for all Borel sets $\mathbb{B} \subset \mathbb{R}$. □

Definition 5.2. Suppose that $\{X_i, i > 0\}$ is sequence of random variables on $\mathbf{\Omega}$ satisfying

1. X_i is adapted to $\mathcal{F}_i$;

2. $\mathbb{E}[|X_i|] < \infty$;

3. $\mathbb{E}[X_i|\mathcal{F}_m] = X_m$ a.s. for all $m < i$.

Then, the sequence $\{X_i, i \geq 0\}$ is said to be a *martingale* with respect to $\{\mathcal{F}_i, i > 0\}$. If 1) and 2) hold, and 3) is replaced by inequality $\mathbb{E}[X_i|\mathcal{F}_m] \geq X_m$ a.s. (or $\mathbb{E}[X_i|\mathcal{F}_m] \leq X_m$ a.s.), then $\{X_i, i \geq 0\}$ is called a *submartingale* (or *supermartingale*).

Furthermore, if 3) is replaced by

3)′ $\mathbb{E}[X_i|\mathcal{F}_m] = 0$ a.s. for all $m < i$.

then X_i is called the martingale difference sequence. □

Example 5.1. Let $X_1, X_2, \cdots$ be independent random variable sequence with $\mathbb{E}[X_i] = 0$. Consider the partial sum process:

$$S_0 = 0, \;\; S_n = \sum_{i=1}^{n} X_i, \;\; n = 1, 2, \cdots$$

Let $\mathcal{F}_n = \sigma(X_m : m < n)$ be the filtration for X_n, i.e. the minimal filtration. Then S_n is a martingale since

$$\begin{aligned}
\mathbb{E}[S_{n+1}|\mathcal{F}_n] &= \mathbb{E}[S_{n+1}|X_1, \cdots, X_n] \\
&= \mathbb{E}[X_n + S_n|X_1, \cdots, X_n] \\
&= \mathbb{E}[X_n|X_1, \cdots, X_n] + S_n \\
&= S_n
\end{aligned}$$

□

5.2 Convergence of Random Variable

An important focus in probability theory is investigating the limiting behavior of a sequence of random variables. This is also known as "large sample theory" or "asymptotic theory". This theory is pivotal in analysing the behaviour of estimators and is termed in that context.

In this section, we aim to introduce some basic definitions and theorems used in the followed chapters which focus on the consistency of estimators. The definitions of different types of convergence are given as follows:

Definition 5.3 (Convergence)**.** Let $X_1, X_2, \cdots$ be a sequence of random variables with cumulative distribution functions $F_1, F_2, \cdots$ and X be some other random variable with cumulative distribution function F.

(1) X_n converges to X in probability if for all $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(|X_n - X| \geq \epsilon\right) = 0$$

This is denoted by $X_n \xrightarrow{P} X$

(2) X_n converges to X almost surely if

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1$$

This is denoted by $X_n \xrightarrow{a.s} X$

(3) X_n converges to X in distribution if

$$\lim_{n\to\infty} F_n(x) = F(x)$$

for every $x \in \mathbb{R}$ at which F is continuous. This is denoted by $X_n \xrightarrow{d} X$.

□

The following theorem shows the relationship between the different types of convergences[95, p. 208]:

Theorem 5.1 (The hierarchy of convergence concepts)**.** Let $X_1, X_2, \cdots$ be a sequence of random variables and X be some other random variable. Then as $n \to \infty$, we have

$$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

□

Next, two of the most notable achievements in probability theory are introduced, i.e. the law of large number(LLN) and central limit theorem(CLT). These two theorems play fundamental role in analysing the limit behaviour of an estimator.

Theorem 5.2 (The strong law of large number)**.** Let $X_1, X_2, \cdots$ be a sequence of independent, identically distributed random variables with common mean μ. Then we have

$$\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{a.s.} \mu$$

□

Theorem 5.3 (The central limit theorem)**.** Let $X_1, X_2, \cdots, X_n$ be a sequence of independent, identically distributed random variables with common mean μ and variance $\mathbb{V}[X_i] = \sigma^2 < \infty$. Then we have

$$\sqrt{n}\frac{\bar{X} - \mu}{\sigma} \xrightarrow{d} \mathcal{N}(0, 1)$$

where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ □

Definition 5.4 (Asymptotic normality)**.** An estimator is said to be asymptotically normal if there exist numbers $\mu_1, \mu_2, \cdots$ and $\sigma_1, \sigma_2, \cdots$ s.t.

$$\frac{\hat{\theta}_n - \mu_n}{\sigma_n} \xrightarrow{d} \mathcal{N}(0, 1)$$

□

Remark 5.1. Normally, $\mu_n = \mathbb{E}[\hat{\theta}_n]$ and $\sigma_n = \sqrt{\mathbb{V}[\hat{\theta}_n]}$. □

Chapter 6

Finding A Position With Distance Dependent Noise — Static Position Case

In target localization problems, the measurement accuracy of the position of a target is normally dependent on the relative distance between the sensor and target. For example, the received power of the radar signal declines as the fourth power of the distance. Therefore the variance of measurement noise should be the functions of the unknown target's position. In practice, this dependency is normally neglected since it makes analysis hard. However, this inaccuracy of the model could lead to the incorrect estimation.

In Section 6.2, the target localization in 1 dimension is formulated. Two cases, namely the robot is static or moving, are described. In the formulation, the measurement noise is assumed to be normally distributed and with 0 *mean and variance dependent on the unknown position of target. Due to this dependency, the classic Maximum Likelihood Estimation(MLE) does not have a closed form solution and the convergence results for the MLE in this problem are largely absent from the literature. In this chapter, we aim to show that the MLE is consistent when the robot is static.*

6.1 Introduction

Consider a typical localization problem in robotic and other engineering fields where a robot or an intelligent agent localizes a static target, for example a bean or a door, via successive "location measurements" of the target using Maximum Likelihood Estimator(MLE). After receiving a new measurement, the robot will estimate the location of the target combining all collected information, it then moves to the newly estimated location of the target. A practical issue is that the location measurement is influenced by the rela-

tive distance between the robot and target, i.e. the smaller the distance between the robot and target, the more accurate the measurement is. One can find many such measurement models in engineering fields, such as radar systems, sensor localization with range measurement[96, 97] and robotics[98]. An important question is whether the robot will eventually arrive at the target's position, i.e. whether the MLE is consistent in this case.

The consistency of the MLE obtained from statistically independent and identically distributed observations has been studied for a long time. Under certain regularity conditions, the MLE is proved to be consistent and asymptotically normal[99–101]. For dependent observations, Bar-shalom[102], Bhat[103] and Crowder [104] show that the MLE is consistent under some mild conditions. An excellent discussion and survey when the observations are not independent and identically distributed is given in [91]. Martingale theory is used in proving consistency of the MLE. Recently, the consistency properties of MLE for statistical models, under the assumption of Hidden Markov models, when parameter value corresponds to a stochastic system observed with noise is studied in [105]. The consistency results of Quasi-MLE with state-dependent were presented by Heijmans[106] and Bollerslev [107]

Though there exist lots of literature related to the consistency and asymptotic behavior of MLE, the consistency analysis for this robotic problem is still absent. The standard conditions are hard to be applied to this kind of problem since the observation is non-i.i.d. and the Markov assumption is unavailable.

6.2 Formulation

Let $\theta_T \in \mathbb{R}^1$ be the true location of the target, $\hat{\theta}_n$ the estimation of target's location at the n-th step and x_n the measurement. It should be noticed that the $\hat{\theta}_n$ is location of robot at $(n+1)$-th step as well. Then the current location of the robot at n-th step, $\hat{\theta}_n,\ n = 1, 2\cdots$, can be described as

$$\begin{cases} \hat{\theta}_n &= \mathbf{S}_n(\hat{\theta}_{0:n-1}, x_{1:n}) \\ x_n &= \theta_T + \omega_n\sqrt{f(\hat{\theta}_{n-1} - \theta_T)} \end{cases} \tag{6.1}$$

where $\hat{\theta}_{0:n-1} = (\hat{\theta}_0, \cdots, \hat{\theta}_{n-1})$, $\mathbf{S}_n(\cdot)$ is the maximum likelihood estimator that used to estimate the location of target and ω_n is independent and identically distributed with zero mean unit variance. And we have following assumptions on function $f(\cdot)$:

C1. $f(\theta)$ is a convex function of $\theta \in \mathbf{\Theta}$ and $\mathbf{\Theta} \subset \mathbb{R}$;

C2. $\min_{\theta \in \mathbf{\Theta}} f(\theta) = f(0) = \sigma^2 > 0$, without loss of generality, we assume that $\sigma^2 = 1$;

C3. $f(\theta)$ is an even function and strictly increasing on $\mathbb{R}^+$ and $f(\infty) = \infty$.

Example 6.1. The typical examples of $f(\theta)$ include polynomial function

$$f(\theta) = |\theta|^\beta + 1, \ \ \beta \geq 1$$

or exponential function

$$f(\theta) = \exp\{|\theta|\}$$

□

Obviously, the expectation and variance of x_n are

$$\mathbb{E}[x_n] = \mathbb{E}\left[\omega_n \sqrt{f(\hat{\theta}_{n-1} - \theta_T)} + \theta_T\right] = \theta_T$$
$$\mathbb{V}[x_n] = \mathbb{V}\left[\omega_n \sqrt{f(\hat{\theta}_{n-1} - \theta_T)} + \theta_T\right] = f(\hat{\theta}_{n-1} - \theta_T)$$

The joint density of $x_{1:n} = (x_1, \cdots, x_n)$ which depends on the parameter $\theta \in \mathbf{\Theta}$ is

$$p(x_{1:n}|\theta) \tag{6.2}$$

the parameter θ is a real-valued constant with unknown true value θ_T, and is estimated by a Borel measurable function $\hat{\theta}_n = \mathbf{S}_n(\hat{\theta}_{0:n-1}, x_{1:n})$ obtained from maximizing the likelihood function

$$L_n \triangleq p(x_{1:n}|\theta) = \prod_{i=1}^{n} p(x_i|x_{1:i-1}, \theta)$$

where

$$p(x_i|x_{1:i-1},\theta) = \frac{L_i}{L_{i-1}}$$

Therefore,

$$\begin{aligned}\hat{\theta}_n &= \arg_{\theta\in\mathbf{\Theta}} \max \left\{ \frac{1}{n}\sum_{i=1}^{n} \log p(x_i|x_{1:i-1},\theta) \right\} \\ &\triangleq \arg_{\theta\in\mathbf{\Theta}} \max \frac{1}{n}\sum_{i=1}^{n} \ell(x_i,\theta,\hat{\theta}_{i-1}) \end{aligned} \tag{6.3}$$

where $\ell(x_i,\theta,\hat{\theta}_{i-1})$ is log-likelihood function at time i.

Since x_i, $i = 1,2,\cdots$, is normally distributed, then the i-th likelihood function is

$$\frac{1}{\sqrt{2\pi f(\hat{\theta}_{i-1}-\theta)}} \exp\left\{ -\frac{(x_i-\theta)^2}{2f(\hat{\theta}_{i-1}-\theta)} \right\}$$

Neglecting the constant term and the log-likelihood function can be written as

$$\ell(x_i,\theta,\hat{\theta}_{i-1}) \triangleq -\log f(\hat{\theta}_{i-1}-\theta) - \frac{(x_i-\theta)^2}{f(\hat{\theta}_{i-1}-\theta)}$$

Then the joint likelihood function is defined by,

$$\frac{1}{n}\sum_{i=1}^{n} \ell(x_i,\theta,\hat{\theta}_{i-1}) = -\frac{1}{n}\sum_{i=1}^{n} \log f(\hat{\theta}_{i-1}-\theta) - \frac{1}{n}\sum_{i=1}^{n} \frac{(x_i-\theta)^2}{f(\hat{\theta}_{i-1}-\theta)} \tag{6.4}$$

There are two cases,

Case I. The robot is static(non-moving), i.e. $\hat{\theta}_i \equiv \hat{\theta}_0, \forall i$, where $\hat{\theta}_0$ is the initial position of robot. It gathers measurements of position of target and performs estimation until it is arbitrarily close to the target. Then from (6.4), the estimation of θ_T via MLE is given by

$$\hat{\theta}_T = \arg\max_{\theta\in\mathbf{\Theta}} \frac{1}{n}\sum_{i=1}^{n} \ell(x_i,\theta,\hat{\theta}_0) \tag{6.5}$$

Case II. The robot will move to the estimated position of θ_T at each step. Then, from

(6.4), estimation via the MLE is given by

$$\hat{\theta}_n = \arg\max_{\theta\in\mathbf{\Theta}} \frac{1}{n}\sum_{i=1}^{n} \ell(x_i, \theta, \hat{\theta}_{i-1}) \tag{6.6}$$

Then $\hat{\theta}_n$ is both the estimated position of target and the next position that the robot will move to.

6.3 Consistency With Distance Dependent Noise

In this section, the robot is assumed to be static, namely Case I described in Section 6.2. We aim to show that the estimation of target's position $\hat{\theta}_T$ defined in (6.5) is consistent. Although the consistency can be established in this case by applying general result, for example see Theorem 17 in [108, P. 117], the approach here reveals the relationship between the MLE, the sample mean and the inverse of variance function f at the sample variance.

Let

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \quad \text{and} \quad s^2 = \frac{1}{n}\sum_{i=1}^{n}(\bar{x} - x_i)^2$$

be the sample mean and (biased) sample variance. Then (6.4) could be rewritten into

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \ell(x_i, \theta, \hat{\theta}_0) \\
&= -\left(\log f(\hat{\theta}_0 - \theta) + \frac{1}{f(\hat{\theta}_0 - \theta)}\frac{1}{n}\sum_{i=1}^{n}(x_i - \theta)^2\right) \\
&= -\left(\log f(\hat{\theta}_0 - \theta) + \frac{1}{f(\hat{\theta}_0 - \theta)}\frac{1}{n}\sum_{i=1}^{n}(x_i - \theta + \bar{x} - \bar{x})^2\right) \\
&= -\left(\log f(\hat{\theta}_0 - \theta) + \frac{1}{f(\hat{\theta}_0 - \theta)}\frac{1}{n}\sum_{i=1}^{n}\left((\bar{x} - \theta)^2 + (x_i - \bar{x})^2 - 2(x_i - \bar{x})(\bar{x} - \theta)\right)\right) && (6.7) \\
&= -\left(\log f(\hat{\theta}_0 - \theta) + \frac{1}{f(\hat{\theta}_0 - \theta)}\left((\bar{x} - \theta)^2 + s^2\right)\right) && (6.8) \\
&= -\left(\log f(\hat{\theta}_0 - \theta) + \frac{s^2}{f(\hat{\theta}_0 - \theta)} + \frac{(\bar{x} - \theta)^2}{f(\hat{\theta}_0 - \theta)}\right) && (6.9) \\
&\triangleq \ell(\theta)
\end{aligned}$$

where (6.8) is due to the cross product terms in (6.7) add to 0. Without loss of generality, we assume that $\hat{\theta}_0 = 0$ and let $s^2 = f(\zeta)$ where $\zeta \in \mathbb{R}^+$. Since $f(\theta)$ is strictly increasing and $f(\infty) = \infty$, then there is a unique root $\zeta > 0$ such that $s^2 = f(\zeta)$.

Then (6.9) becomes

$$-\left(\log f(\theta) - \frac{s^2}{f(\theta)} - \frac{(\bar{x}-\theta)^2}{f(\theta)}\right) \triangleq -\left(\log f(\theta) - \frac{f(\zeta)}{f(\theta)} - \frac{(\bar{x}-\theta)^2}{f(\theta)}\right) \tag{6.10}$$

And the score function is

$$\begin{aligned} D(\theta) \triangleq & \frac{\partial}{\partial\theta}\left\{-\left(\log f(\theta) + \frac{f(\zeta)}{f(\theta)} + \frac{(\bar{x}-\theta)^2}{f(\theta)}\right)\right\} \\ = & \frac{1}{f(\theta)^2}\left(f'(\theta)f(\zeta) + f'(\theta)(\theta-\bar{x})^2 - f'(\theta)f(\theta) - 2f(\theta)\,(\theta-\bar{x})\right) \end{aligned} \tag{6.11}$$

Following propositions will be useful.

Proposition 6.1.

$$\{-\zeta, \zeta\} = \arg\max_{\theta\in\mathbb{R}}\left\{-\log f(\theta) - \frac{f(\zeta)}{f(\theta)}\right\} \tag{6.12}$$

□

Proof. Since the function to be maximized is symmetric. Thus it is sufficient to discuss the case $\theta \in \mathbb{R}^+$.

For $\theta \in \mathbb{R}^+$, since $f(\theta)$ is increasing in $\mathbb{R}^+$, we have

$$\begin{aligned} & \frac{\partial}{\partial\theta}\left\{-\log f(\theta) - \frac{f(\zeta)}{f(\theta)}\right\} \\ = & \frac{f'(\theta)}{f^2(\theta)}\left(f(\zeta) - f(\theta)\right) \begin{cases} > 0 & \text{if } \theta < \zeta \\ < 0 & \text{if } \theta > \zeta \\ = 0 & \text{if } \theta = \zeta \end{cases} \end{aligned} \tag{6.13}$$

Then we know that

$$\zeta = \arg\max_{\theta\in\mathbb{R}^+}\left\{-\log f(\theta) - \frac{f(\zeta)}{f(\theta)}\right\} \tag{6.14}$$

Therefore,

$$\{-\zeta, \zeta\} = \arg\max_{\theta \in \mathbb{R}} \left\{ -\log f(\theta) - \frac{f(\zeta)}{f(\theta)} \right\} \tag{6.15}$$

□

Proposition 6.2. If $\bar{x} > 0$, $-\frac{(\theta-\bar{x})^2}{f(\theta)}$ is strictly increasing for $\theta \in [0, \bar{x}]$, and

$$\bar{x} = \arg\max_{\theta \in \mathbb{R}} \left\{ -\frac{(\theta - \bar{x})^2}{f(\theta)} \right\}.$$

□

Proof. Let $\theta \in [0, \bar{x}]$, then

$$\frac{\partial}{\partial \theta} \left\{ -\frac{(\theta - \bar{x})^2}{f(\theta)} \right\} = -\frac{(\theta - \bar{x}) \left((\theta - \bar{x}) f'(\theta) - 2f(\theta) \right)}{f(\theta)^2} > 0$$

This implies that

$$-\frac{(\theta - \bar{x})^2}{f(\theta)}$$

is strictly increasing when $\theta \in [0, \bar{x}]$.

Since $-\frac{(\theta-\bar{x})^2}{f(\theta)} \leq 0$ and attains 0 if and only if $\theta = \bar{x}$, thus

$$\bar{x} = \arg\max_{\theta \in \mathbb{R}} \left\{ -\frac{(\theta - \bar{x})^2}{f(\theta)} \right\}.$$

□

Proposition 6.3. Suppose $\hat{\theta}_T$ is defined in (6.5). $\hat{\theta}_T > 0$ if $\bar{x} > 0$ and $\hat{\theta}_T < 0$ if $\bar{x} < 0$. □

Proof. Suppose $\bar{x} > 0$, then we have

$$-\frac{(\theta - \bar{x})^2}{f(\theta)}\Big|_{|\theta|} \geq -\frac{(\theta - \bar{x})^2}{f(\theta)}\Big|_{-|\theta|} \tag{6.16}$$

equality only holds at $\theta = 0$.

From proposition 6.1, we have

$$\begin{aligned} &\left(-\log f(\theta) - \frac{f(\zeta)}{f(\theta)} - \frac{(\theta-\bar{x})^2}{f(\theta)}\right)\Bigg|_{|\theta|} \\ \geq &\left(-\log f(\theta) - \frac{f(\zeta)}{f(\theta)} - \frac{(\theta-\bar{x})^2}{f(\theta)}\right)\Bigg|_{-|\theta|} \end{aligned} \tag{6.17}$$

equality only holds at $\theta = 0$. Then $\hat{\theta}_T > 0$.

Similarly, if $\bar{x} < 0$, then $\hat{\theta}_T < 0$. □

Proposition 6.3 states that it is sufficient to discuss the consistency of MLE with the assumption $\bar{x} > 0$. If $\bar{x} < 0$, the analysis is similar.

Theorem 6.1 and Lemma 6.1 are need in proving Theorem 6.2.

Theorem 6.1. [109] Suppose function h is continuously differentiable at a. If $h'(a) < 0$, there exists a neighbourhood $\mathcal{B}_\mu(a)$ such that $\forall t \in \mathcal{B}_\mu(a)$,

$$\begin{aligned} t < a &\Rightarrow h(t) > h(a) \\ t > a &\Rightarrow h(t) < h(a) \end{aligned}$$

□

Recall $s^2 = \frac{1}{n}\sum_{i=1}^{n}(\bar{x} - x_i)^2 = f(\zeta)$, then we have following lemma

Lemma 6.1. $s^2 \xrightarrow{a.s.} f(\theta_T)$ □

Proof. By Theorem 5.2,

$$\begin{aligned} s^2 &= \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \\ &\xrightarrow{a.s.} \mathbb{E}[x_1^2]) - (\mathbb{E}[x_1])^2 \\ &= f(\theta_T) \end{aligned}$$

□

Now we prove the key theorem, Theorem 6.2, of this section.

Theorem 6.2. Suppose $\hat{\theta}_T$ is the MLE, then $\hat{\theta}_T$ satisfies

$$\hat{\theta}_T \in [(\text{sgn}(\bar{x}) \cdot \zeta) \wedge \bar{x}, (\text{sgn}(\bar{x}) \cdot \zeta) \vee \bar{x}] \quad \textit{eventually}, \quad \textit{a.s.} \tag{6.18}$$

where sgn is sign function. □

Proof. Firstly, assume that $\bar{x} > 0$. Then from Proposition 6.3, we have $\hat{\theta}_n > 0$, therefore we only need to consider the behaviour of likelihood function defined in (6.5) when $\theta \in \mathbb{R}^+$. With these assumptions, we aim to prove

$$\hat{\theta}_T \in [\zeta \wedge \bar{x}, \zeta \vee \bar{x}] \quad \textit{eventually}, \quad \textit{a.s.} \tag{6.19}$$

the proof can be separated into two parts.

Part I: $0 < \zeta < \bar{x}, \theta \in \mathbb{R}^+$

Assume that $0 < \zeta < \bar{x}$. Then $\forall \theta < \zeta$ and $\theta > 0$, from Proposition 6.1 and 6.2, we know

$$-\log f(\theta) - \frac{f(\zeta)}{f(\theta)} \quad \text{and} \quad -\frac{(\theta - \bar{x})^2}{f(\theta)}$$

are both strictly increasing. Therefore $\frac{1}{n}\ell(x_i, \theta)$ is strictly increasing $\forall \theta < \zeta$ and we must have $\hat{\theta}_T \geq \zeta$.

On the other hand, from Proposition 6.2 and from (6.13), $\bar{x} = \arg\max_{\theta \in \mathbb{R}} \left\{ -\frac{(\theta - \bar{x})^2}{f(\theta)} \right\}$ and $-\log f(\theta) - \frac{f(\zeta)}{f(\theta)}$ is strictly decreasing $\forall \theta > \bar{x} > \zeta$, then

$$\begin{aligned}
\max_{\forall \theta \geq \bar{x}} \ell(\theta) &= \max_{\forall \theta \geq \bar{x}} \left\{ -\log f(\theta) - \frac{f(\zeta)}{f(\theta)} - \frac{(\theta - \bar{x})^2}{f(\theta)} \right\} \\
&\leq \max_{\forall \theta \geq \bar{x}} \left\{ -\log f(\theta) - \frac{f(\zeta)}{f(\theta)} \right\} + \max_{\forall \theta \geq \bar{x}} \left\{ -\frac{(\theta - \bar{x})^2}{f(\theta)} \right\} \\
&= \left\{ -\log f(\theta) - \frac{f(\zeta)}{f(\theta)} \right\} \Big|_{\theta = \bar{x}} + \left\{ -\frac{(\theta - \bar{x})^2}{f(\theta)} \right\} \Big|_{\theta = \bar{x}} \\
\implies \bar{x} = \arg\max_{\forall \theta \geq \bar{x}} \ \ell(\theta) &\implies \hat{\theta}_T \leq \bar{x}
\end{aligned} \tag{6.20}$$

Therefore $\hat{\theta}_T \in [\zeta, \bar{x}]$.

Part II: $\zeta > \bar{x} > 0, \theta \in \mathbb{R}^+$

Assume $\zeta > \bar{x} > 0$. From Proposition 6.1 and 6.2, we know that $\forall \theta < \bar{x}$ and $\theta > 0$, $\ell(\theta)$ is increasing function, thus $\hat{\theta}_T > \bar{x}$.

Consider

$$\begin{aligned} \operatorname{sgn}(D(\zeta)) =& \operatorname{sgn}\left(f'(\zeta)f(\zeta) + f'(\zeta)(\zeta-\bar{x})^2 - f'(\zeta)f(\zeta) - 2f(\zeta)\,(\zeta-\bar{x})\right) && (6.21)\\ =& \operatorname{sgn}\left(f'(\zeta)(\zeta-\bar{x})^2 - 2f(\zeta)\,(\zeta-\bar{x})\right) && (6.22)\end{aligned}$$

By Law of Theorem 5.2 and Lemma 6.1,

$$\bar{x} \xrightarrow{a.s.} \theta_T \;\text{ and }\; s^2 \xrightarrow{a.s.} f(\theta_T) \tag{6.23}$$

then

$$\zeta = (f)^{-1}(s^2) \xrightarrow{a.s.} |\theta_T| \tag{6.24}$$

Thus, for any arbitrary small ϵ, there exists (random) $N(\epsilon) \in \mathbb{N}$, for $n > N(\epsilon)$, we have

$$\zeta - \bar{x} < \epsilon \quad a.s. \tag{6.25}$$

Therefore

$$\frac{2f(\zeta)}{f'(\theta)} > \zeta - \bar{x} \quad a.s. \Rightarrow \operatorname{sgn}(D(\zeta)) = -1 \;\; a.s. \tag{6.26}$$

On the other hand, rewrite

$$\begin{aligned} \ell(\theta) =& -\log f(\theta) - \frac{f(\zeta)}{f(\theta)} - \frac{(\theta-\bar{x})^2}{f(\theta)}\\ =& \underbrace{-\log f(\theta) - \frac{f(\zeta)}{f(\theta)} - \frac{(\theta-\zeta)^2}{f(\theta)}}_{\triangleq \ell_1(\theta)} - \frac{(\zeta-\bar{x})^2}{f(\theta)} - 2\frac{(\zeta-\bar{x})(\theta-\zeta)}{f(\theta)}\\ =& \ell_1(\theta) + (\zeta-\bar{x})\underbrace{\left(-\frac{(\zeta-\bar{x})}{f(\theta)} - 2\frac{(\theta-\zeta)}{f(\theta)}\right)}_{\triangleq \ell_2(\theta)}\end{aligned}$$

$$=\ell_1(\theta) + (\zeta - \bar{x})\ell_2(\theta) \tag{6.27}$$

Then $\forall \theta \in \mathbb{R}^+ \setminus \{\zeta\}$

$$\begin{aligned}
\ell_1(\zeta) - \ell_1(\theta) =& -\log f(\zeta) - \frac{f(\zeta)}{f(\zeta)} - \frac{(\zeta-\zeta)^2}{f(\zeta)} + \\
& \log f(\theta) + \frac{f(\zeta)}{f(\theta)} + \frac{(\theta-\zeta)^2}{f(\theta)} \\
=& -\log f(\zeta) - 1 + \log f(\theta) + \frac{f(\zeta)}{f(\theta)} + \frac{(\theta-\zeta)^2}{f(\theta)} \\
=& -\log \frac{f(\zeta)}{f(\theta)} + \frac{f(\zeta)}{f(\theta)} - 1 + \frac{(\theta-\zeta)^2}{f(\theta)} > 0
\end{aligned} \tag{6.28}$$

Thus

$$\zeta = \arg\max_{\theta \in \mathbb{R}^+} \ell_1(\theta) \tag{6.29}$$

Compute

$$\begin{aligned}
\frac{\partial}{\partial\theta}\ell_2(\theta)\Big|_{\theta=\zeta} =& \frac{\partial}{\partial\theta}\left\{-\frac{(\zeta-\bar{x})}{f(\theta)} - 2\frac{(\theta-\zeta)}{f(\theta)}\right\}\Big|_{\theta=\zeta} \\
=& \frac{1}{f^2(\theta)}\left(-2f(\theta) + f'(\theta)(2\theta - \zeta - \bar{x})\right)\Big|_{\theta=\zeta} \\
=& \frac{1}{f^2(\zeta)}\left(-2f(\zeta) + f'(\zeta)(\zeta - \bar{x})\right) \\
<& 0 \quad a.s.
\end{aligned} \tag{6.30}$$

Inequality (6.30) is because of

$$\frac{1}{f(\theta)} = 0 \quad \text{and} \quad f'(\theta)\Big|_{\theta=0} = 0 \tag{6.31}$$

and $\zeta < \infty$ $a.s.$ and ζ is also bounded away from 0, thus from (6.25), we have $\frac{\partial}{\partial\theta}\{\ell_2(\theta)\}|_{\theta=\zeta}$ is a negative value which is bounded away from 0.

From $\frac{\partial}{\partial\theta}\ell_2(\theta)|_{\theta=\zeta} < 0$ $a.s.$ and Theorem 6.1, we know that there exists a neighbour-

hood $\mathcal{B}_\mu(\zeta)$ of ζ such that, $\forall\theta \in (\zeta, \zeta+\mu)$,

$$\ell_2(\zeta) > \ell_2(\theta) \tag{6.32}$$

Since $\frac{\partial}{\partial\theta}\ell_2(\theta)|_{\theta=\zeta}$ is a negative value which is bounded away from 0, therefore we always can find such a positive constant μ.

Then for $\theta \in [\zeta, \zeta+\mu)$,

$$\begin{cases} \zeta = & \arg\max_{\theta\in[\zeta,\zeta+\mu)} \ \ell_1(\theta) \ \text{, and} \\ \zeta = & \arg\max_{\theta\in[\zeta,\zeta+\mu)} \ \{(\zeta-\bar{x})\ell_2(\theta)\} \end{cases} \tag{6.33}$$

$$\Longrightarrow \zeta = \arg\max_{\theta\in[\zeta,\zeta+\mu)} \{\ell_1(\theta) + (\zeta-\bar{x})\ell_2(\theta)\} \tag{6.34}$$

$$\Longleftrightarrow \zeta = \arg\max_{\theta\in[\zeta,\zeta+\mu)} \ \ell(\theta) \tag{6.35}$$

On the other hand, since

$$\zeta = \arg\max_{\theta\geq\zeta} \ \ell_1(\theta) \tag{6.36}$$

then $\forall\theta \geq (\zeta+\mu)$, there exists a positive value η,

$$\ell_1(\zeta) - \max_{\theta\geq(\zeta+\mu)} \ell_1(\theta) > \eta \tag{6.37}$$

Thus $\forall\theta \geq (\zeta+\mu)$

$$\begin{aligned} &\ell(\zeta) - \ell(\theta) \\ =&\ell_1(\zeta) - \frac{(\zeta-\bar{x})^2}{f(\zeta)} - \left(\ell_1(\zeta) - \frac{(\zeta-\bar{x})^2}{f(\theta)} - 2(\zeta-\bar{x})\frac{\theta-\zeta}{f(\theta)}\right) \\ >&\ell_1(\zeta) - \ell_1(\theta) - \frac{(\zeta-\bar{x})^2}{f(\zeta)} \\ >&\eta - \frac{(\zeta-\bar{x})^2}{f(\zeta)} > 0 \quad a.s. \end{aligned} \tag{6.38}$$

Thus from

$$\begin{cases} \zeta = \arg\max_{\theta\in[\zeta,\zeta+\mu)} \ \ell(\theta) \\ \ell(\zeta) > \max_{\theta\geq\zeta+\mu} \ \ell(\theta) \ \ a.s. \end{cases} \tag{6.39}$$

we have

$$\zeta = \arg\max_{\theta\geq\zeta} \ \ell(\theta) \quad a.s. \tag{6.40}$$

Therefore, combining $\hat{\theta}_T > \bar{x}$, $D(\zeta) < 0$ and $\zeta = \arg\max_{\theta\geq\zeta} \ \ell(\theta)$, we have $\hat{\theta}_T \in [\bar{x}, \zeta], \ \ a.s.$.

From **Part I** and **Part II**,

$$\hat{\theta}_T \in [\zeta \wedge \bar{x}, \zeta \vee \bar{x}] \quad eventually, \ \ a.s. \tag{6.41}$$

Similarly, if $\bar{x} < 0$, we have

$$\hat{\theta}_T \in [-\zeta \wedge \bar{x}, -\zeta \vee \bar{x}] \quad eventually, \ \ a.s. \tag{6.42}$$

In conclusion,

$$\hat{\theta}_T \in [(\mathrm{sgn}(\bar{x}) \cdot \zeta) \wedge \bar{x}, (\mathrm{sgn}(\bar{x}) \cdot \zeta) \vee \bar{x}] \quad eventually, \ \ a.s. \tag{6.43}$$

□

Remark 6.1. Two examples, Fig.6.1 and 6.2, are used to illustrate the interval of $\hat{\theta}_T$ given in the Theorem 6.2 with different sample size. In these two example, the used variance function is $f(\theta) = \theta^2 + 1$ and $\theta_T = 20$.

On the other hand, it should be noticed that there may exist multi-modal(at least one) within $[(\mathrm{sgn}(\bar{x}) \cdot \zeta) \wedge \bar{x}, (\mathrm{sgn}(\bar{x}) \cdot \zeta) \vee \bar{x}]$, however, the one which maximize the function over $\boldsymbol{\Theta}$ is the required MLE. □

Based on the above propositions and theorem, we now give the main theorem 6.3

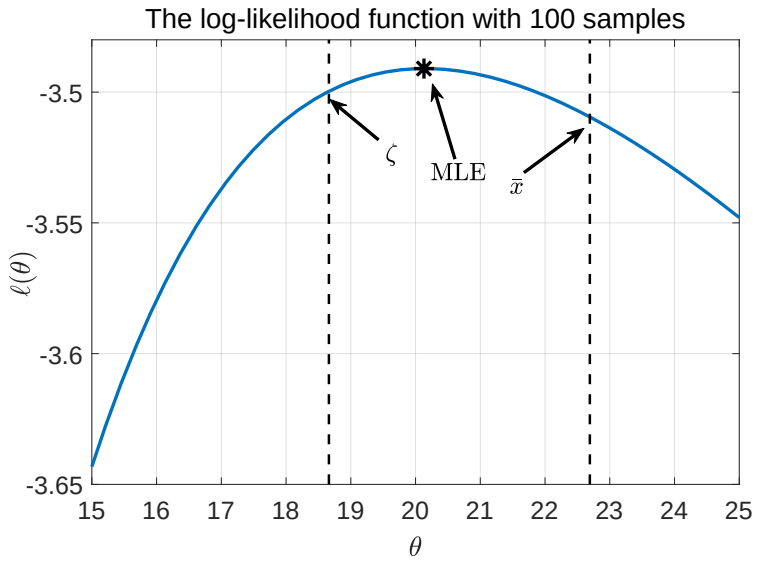


Figure 6.1: The illustration of log-likelihood function $\ell(\theta)$ with 100 samples.

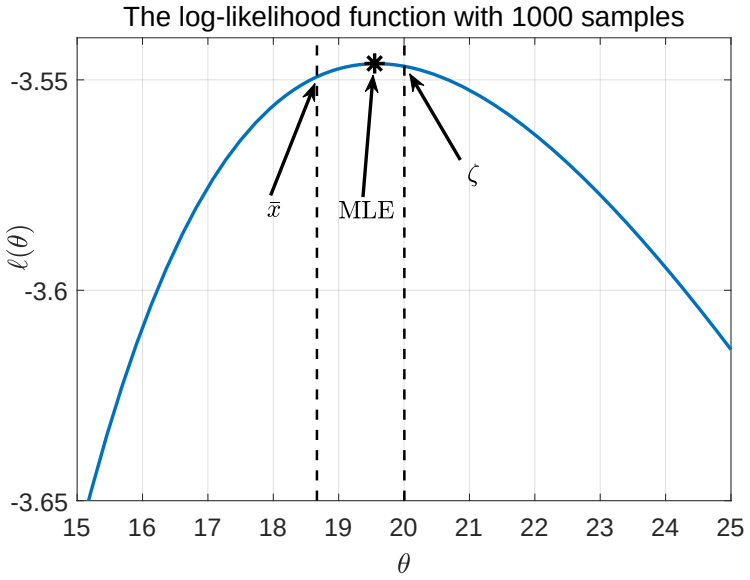


Figure 6.2: The illustration of log-likelihood function $\ell(\theta)$ with 1000 samples.

Theorem 6.3. Suppose $\hat{\theta}_T$ is the MLE, then

$$\hat{\theta}_T \xrightarrow{a.s.} \theta_T \tag{6.44}$$

□

Proof. From Theorem 5.2 and Lemma 6.1, we have

$$\bar{x} \xrightarrow{a.s.} \theta_T \text{ and } s^2 \xrightarrow{a.s.} f^2(\theta_T) \tag{6.45}$$

then from Theorem 6.2,

$$\hat{\theta}_T \xrightarrow{a.s.} \theta_T \tag{6.46}$$

□

6.4 Conclusions

In this chapter, finding location via distance dependent noise using MLE is formulated. Two cases, the robot is non-moving or moving, are described as well. A natural question is if the MLE is a consistent estimator in these two cases. In this chapter, by analyzing the likelihood function and describing the inverse of variance function f at MLE in relation to two natural start point for estimation, sample variance and sample mean, we derive the consistency result for the first case in Section 6.3. The more complicated case that the robot is assumed to be moveable will be considered in the next chapter.

Chapter 7

Finding A Position With Distance Dependent Noise — Dynamical Position Case

In Chapter 6, the problem of finding a position with distance dependent noise was formulated and the case I that the robot is static, was analysed. The MLE in case I was shown to be a consistent estimator. Another question is how the MLE behaves when the robot moves to the estimated location of target after each step, i.e. case II. In this case, the two main challenges are: 1) the joint log-likelihood function $\sum_{i=1}^{n} \ell(\theta, x_i, \hat{\theta}_{i-1})$ is complicated and, 2) the estimate $\hat{\theta}_n$ is dependent on all previous measurements $x_1, \cdots, x_n$ as shown in (6.3). As a result, the conventional consistency theorems are inapplicable due to the non-Markovian dependency. In this chapter, we aim to prove that the MLE for case II is consistent under some conditions.

7.1 Introduction and General Consistency Theorem

In this chapter, a more complex robotic problem is considered. In Chapter 6, the robot was assumed to be static and collected measurements of the position of the target. However, if the robot moves to the estimated location of target after each step, the log-likelihood function becomes, as shown in Chapter 6.1,

$$\frac{1}{n}\sum_{i=1}^{n} \ell(x_i, \theta, \hat{\theta}_{i-1}) = \frac{1}{n}\sum_{i=1}^{n} \left(\log \frac{1}{f\left(\hat{\theta}_{i-1} - \theta\right)} - \frac{(x_i - \theta)^2}{f\left(\hat{\theta}_{i-1} - \theta\right)} \right) \tag{7.1}$$

And the estimate at n-th step, $\hat{\theta}_n$, is given by

$$\hat{\theta}_n = \arg\max_{\theta \in \mathbf{\Theta}} \frac{1}{n}\sum_{i=1}^{n} \ell(x_i, \theta, \hat{\theta}_{i-1}) \tag{7.2}$$

Obviously, $\hat{\theta}_n$ is a function of $x_1, \cdots, x_n$. In other words, the random variable $\hat{\theta}_n$ is dependent on the all previous information. As a result, conventional consistency theorems are inapplicable since the independence assumption or Markov condition are not satisfied. Fortunately, martingale theory is shown to be a powerful tool in dealing with the dependent random variables as the conditional expectation is involved. In the following sections, we aim to investigate the behaviour of the MLE given in (7.2) using martingale theory.

Two definitions and the general consistency theorem, Theorem 7.1, for MLE using martingale theory are introduced.

Definition 7.1 (*Uniformly Strong Law of Large Number for Martingale,USLLNM*)**.** Let $\mathcal{Y} \times \mathbf{\Theta}$ be the Cartesian product of $\mathcal{Y} \subset \mathbb{R}$ and $\mathbf{\Theta} \subset \mathbb{R}$. Let $q_i(y_i, \theta)$ be a real-valued function defined on $\mathcal{Y} \times \mathbf{\Theta}$. Assume that $q_i(\cdot, \theta)$ is measurable for every $\theta \in \mathbf{\Theta}$ and $y_1, \cdots, y_n$ is a sequence of random variables on $\mathcal{Y}$. Let $\mathcal{F}_n$ be the σ-field generated by $y_1, \cdots, y_n$. Then $q_i(y_i, \theta)$ is said to satisfiy the Uniform Strong Law of Large Number for Martingale(USLLNM) if

$$\sup_{\theta \in \mathbf{\Theta}} \left| \frac{1}{n} \sum_{i=1}^{n} \Big(q_i(y_i, \theta) - \mathbb{E}\left[q_i(y_i, \theta) | \mathcal{F}_{i-1} \right] \Big) \right| \xrightarrow{a.s.} 0$$

□

Definition 7.2 (*Identifiably Unique Maximizer*)**.** Let $q_i(\theta)$ be a real valued continuous function on a compact space $\mathbf{\Theta} \subset \mathbb{R}$ such that $q_i(\theta)$ has a maximum at θ_i^*, $i = 1, 2, \cdots, n$. Let $\mathcal{B}_\epsilon(\theta_i^*)$ be an open ball centered at θ_i^* with fixed radius $\epsilon > 0$. Define the neighbourhood $\eta_i(\epsilon) = \mathcal{B}_\epsilon(\theta_i^*) \cap \mathbf{\Theta}$. The maximizer θ_i^* is said to be identifiably unique if and only if for any $\epsilon > 0$

$$\limsup_{n \to \infty} \left\{ \max_{\theta \in \bar{\eta}(\epsilon)} q_n(\theta) - q_n(\theta_n^*) \right\} < 0$$

□

Next, we give a general consistency theorem 7.1. This is the martingale version of the theorem in [110] and [111].

Theorem 7.1. Assumptions:

A1. θ is in the interior of $\mathbf{\Theta} \subset \mathbb{R}^1$ and $\mathbf{\Theta}$ is compact;

A2. $\left\{\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1})\right\}$ satisfies USLLNM, i.e.

$$\sup_{\theta \in \mathbf{\Theta}} \Big| \frac{1}{n} \sum_{i=1}^{n} \Big(\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) - \mathbb{E}_{\theta_T} \left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) | \mathcal{F}_{i-1} \right] \Big) \Big| \xrightarrow{a.s.} 0$$

where $\mathcal{F}_{i-1} = \sigma(x_1, \cdots, x_i)$.

A3. $\forall \theta \in \mathbf{\Theta}$, θ_T is the Identifiably Unique Maximizer(IUM) of

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \big| \mathcal{F}_{i-1} \right].$$

If A1., A2., A3. are satisfied, then the MLE is (weakly) consistent, i.e. $\hat{\theta}_n \xrightarrow{p} \theta_T$ under P_{θ_T}. □

Proof. Let $\mathcal{B}_\epsilon(\theta_T)$ be the open ball with radius ϵ and center θ_T, i.e. $\mathcal{B}_\epsilon(\theta_T) = \{\theta_T \in \mathbf{\Theta} : |\theta - \theta_T| < \epsilon\}$. We need to show that $P_{\theta_T}(\hat{\theta}_n \notin \mathcal{B}_\epsilon(\theta_T)) \to 0$.

Since θ_T is the IUM of

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \big| \mathcal{F}_{i-1} \right]$$

thus, according to the definition of IUM, we have

$$\eta \triangleq \sup_{\theta \in \mathbf{\Theta} \setminus \mathcal{B}_\epsilon(\theta_T)} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \big| \mathcal{F}_{i-1} \right] < 0 \tag{7.3}$$

Note that $\hat{\theta}_n \notin \mathcal{B}_\epsilon(\theta_T)$ implies

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \big| \mathcal{F}_{i-1} \right] \\ \leq & \sup_{\theta \in \mathbf{\Theta} \setminus \mathcal{B}_\epsilon(\theta_T)} \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \big| \mathcal{F}_{i-1} \right] \\ = & \eta \end{aligned} \tag{7.4}$$

Let

$$E_n^1 = \left\{ \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1} \right] \leq \eta \right\}$$

thus

$$P(\hat{\theta}_n \notin \mathcal{B}_\epsilon(\theta_T)) \leq P(E_n^1)$$

Consider the following events

$$\begin{aligned} E_n^2 = \Bigg\{ & \left| \frac{1}{n} \sum_{i=1}^{n} \left(\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right) - \right. \\ & \left. \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1} \right] \right| > \frac{|\eta|}{2} \Bigg\} \\ E_n^3 = \Bigg\{ & \sup_{\theta \in \Theta} \left| \frac{1}{n} \sum_{i=1}^{n} \left(\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right) - \right. \\ & \left. \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1} \right] \right| > \frac{|\eta|}{2} \Bigg\} \end{aligned}$$

Note that $E_n^2 \subseteq E_n^3$, so $P(E_n^2) \leq P(E_n^3)$.

Consider the complementary event $\bar{E}_n^2$ of E_n^2, we have

$$\begin{aligned} & \left| \frac{1}{n} \sum_{i=1}^{n} \left(\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right) - \right. \\ & \left. \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1} \right] \right| \leq \frac{|\eta|}{2} \end{aligned} \tag{7.5}$$

Note

$$\frac{1}{n} \sum_{i=1}^{n} \mathbb{E}_{\theta_T} \left[\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1} \right] \leq \eta < 0$$

Then we must have

$$\frac{1}{n} \sum_{i=1}^{n} \left(\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right) < \frac{\eta}{2}$$

otherwise, (7.5) will not hold.

However, since $\hat{\theta}_n$ is the Maximizer of log-likelihood function over $\mathbf{\Theta}$, i.e.

$$\hat{\theta}_n = \max_{\theta \in \mathbf{\Theta}} \frac{1}{n} \sum_{i=1}^{n} \ell(x_i, \theta, \hat{\theta}_{i-1})$$

then

$$\frac{1}{n} \sum_{i=1}^{n} \left(\ell(x_i, \hat{\theta}_n, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right) \geq 0$$

Therefore we have the conclusion that $E_n^1 \cap \bar{E}_n^2 = \varnothing$.

Therefore,

$$\begin{aligned} P(E_n^1) &\leq P(E_n^1 \cup E_n^2) \\ &\leq P(E_n^1 \cap \bar{E}_n^2) + P(E_n^2) \\ &= P(\varnothing) + P(E_n^2) \\ &\leq P(E_n^3) \end{aligned}$$

Since $\left\{ \ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right\}$ satisfies the USLLNM, we have $P(E_n^3) \xrightarrow{a.s.} 0$ and then

$$\hat{\theta}_n \xrightarrow{p} \theta_T \quad \text{under} \quad P_{\theta_T}$$

□

7.2 Consistency With Distance Dependent Noise

In next two sections, we aim to show that the likelihood function (7.1) satisfies A1, A2 and A3, and therefore the MLE is consistent. The assumption A1 is easily to be satisfied as we could assume that $\mathbf{\Theta}$ to be a compact set.

7.2.1 Assumption A2

Assumption A2 requires that $\left\{ \ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \right\}$ satisfies USLLNM. The main theorem is given in Theorem 7.3 which follows Theorem 7.2, Lemma 7.2 and Lemma 7.1.

Theorem 7.2. [112] Let $Y_n = y_1 + \cdots + y_n$ be a martingale with $\mathbb{E}[y_n] = 0$. If

$$\sum_{i=1}^{\infty} \frac{\mathbb{E}[|y_i|^{2\alpha}]}{i^{1+\alpha}} < \infty$$

then

$$\lim_{n\to\infty} \frac{Y_n}{n} = 0 \quad \text{a.s.}$$

□

Then, the Lemma 7.1 is straightforward from Theorem 7.2,

Lemma 7.1 (Strong Law of Large Numebr for Martingale, SLLNM). Let $Y_n = y_1 + \cdots + y_n$ be a martingale with $\mathbb{E}[y_n] = 0$. If

$$\sup_i \mathbb{E}[|y_i|^2] < \infty$$

then

$$\lim_{n\to\infty} \frac{Y_n}{n} = 0 \quad \text{a.s.}$$

□

We also have following lemma.

Lemma 7.2. Under P_{θ_T},

$$\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) - \mathbb{E}_{\theta_T}\left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1}\right]$$

is martingale difference. □

Proof. Under P_{θ_T}, we can write

$$x_i = \theta_T + \omega_i \sqrt{f(\hat{\theta}_{i-1} - \theta)} \tag{7.6}$$

where ω_i is i.i.d standard normal random variable.

Then $\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1})$ can be written into

$$\begin{aligned}
&\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \\
=& -\frac{(\theta - \theta_T)^2}{f(\hat{\theta}_{i-1} - \theta)} + \left(1 - \frac{f(\hat{\theta}_{i-1} - \theta_T)}{f(\hat{\theta}_{i-1} - \theta)}\right) \omega_i^2 + 2\frac{f(\hat{\theta}_{i-1} - \theta_T)\omega_i}{f(\hat{\theta}_{i-1} - \theta)}
\end{aligned} \tag{7.7}$$

so

$$\mathbb{E}_{\theta_T}\left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1})\right] = -\frac{(\theta - \theta_T)^2}{f(\hat{\theta}_{i-1} - \theta_T)} + \left(1 - \frac{f(\hat{\theta}_{i-1} - \theta_T)}{f(\hat{\theta}_{i-1} - \theta)}\right) \tag{7.8}$$

Since $\hat{\theta}_{i-1}$ is $\mathcal{F}_{i-1}$ measurable and ω_i is independent of $\mathcal{F}_{i-1}$ with 0 mean and variance 1. Hence,

$$\begin{aligned}
&\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) - \mathbb{E}_{\theta_T}\left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1})\right] \\
=& \left(1 - \frac{f(\hat{\theta}_{i-1} - \theta_T)}{f(\hat{\theta}_{i-1} - \theta)}\right)(\omega_i^2 - 1) + 2\frac{f(\hat{\theta}_{i-1} - \theta_T)}{f(\hat{\theta}_{i-1} - \theta)}\omega_i
\end{aligned} \tag{7.9}$$

is martingale difference process. □

Then we have following USLLNM theorem.

Theorem 7.3 (Uniformly Strong Law of Large Number for Martingale, USLLNM)**.** Let $\{q_i(y_i, \theta), i \geq 1\}$ be a Borel measurable function on $\boldsymbol{\mathcal{Y}} \times \boldsymbol{\Theta}$ and $\mathcal{F}_i$ be an increasing sequence of $\sigma-$fields with $\{y_i, i \geq 1\}$, where $\mathcal{Y}$ is a Borel set such that $P(y_i \in \boldsymbol{\mathcal{Y}}) = 1$ and $\boldsymbol{\Theta}$ is the compact parameter space of θ such that $q_i(y_i, \theta)$ is a continuous function. If the following two assumptions are satisfied, $\forall \theta \in \boldsymbol{\Theta}$

1. $\sup_i \mathbb{E}\left[|q_i(y_i, \theta)|^2\right] < \infty$;
2. $q_i(y_i, \theta)$ is differentiable a.s. $\forall i$.

Then

$$\sup_{\theta \in \boldsymbol{\Theta}} \left| \frac{1}{n} \sum_{i=1}^{n} \left(q_i(y_i, \theta) - \mathbb{E}\left[q_i(y_i, \theta) | \mathcal{F}_{i-1}\right] \right) \right| \xrightarrow{a.s.} 0 \tag{7.10}$$

□

Proof. By Theorem 2 in [113], we know that (7.10) holds if

1). $\boldsymbol{\Theta}$ is compact;

2). $\frac{1}{n}\sum_{i=1}^{n}\left(q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right)\xrightarrow{a.s.}0, \forall\theta\in\boldsymbol{\Theta}$;

3). $\frac{1}{n}\sum_{i=1}^{n}\left(q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right)$ is Strongly Stochastically Equicontinuous(S.S.E).

are satisfied.

Since

$$\mathbb{E}\ \left[q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\,\mathcal{F}_{i-1}\right]|\,\mathcal{F}_{i-1}\right]=0$$

then

$$q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]$$

is a martingale difference process.

By Hölder's inequality and the conditional Jensen's inequality,

$$\begin{aligned}
&\mathbb{E}\left[\left|q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right|^2\right]\\
\leq&\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2\right]+\mathbb{E}\left[\left|\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right|^2\right]+2\mathbb{E}\left[\ \left|q_i(y_i,\theta)\right|\cdot\left|\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right|\ \right] &(7.11)\\
\leq&\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2\right]+\mathbb{E}\left[\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2|\mathcal{F}_{i-1}\right]\right]+\\
&2\left(\mathbb{E}\left[\ \left|q_i(y_i,\theta)\right|^2\ \right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\left|\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right|^2\right]\right)^{\frac{1}{2}} &(7.12)\\
\leq&2\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2\right]+2\left(\mathbb{E}\left[\ \left|q_i(y_i,\theta)\right|^2\ \right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2|\mathcal{F}_{i-1}\right]\right]\right)^{\frac{1}{2}} &(7.13)\\
\leq&4\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2\right] &(7.14)
\end{aligned}$$

Then

$$\sup_i\mathbb{E}\left[\left|q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right|^2\right]\leq 4\sup_i\mathbb{E}\left[\left|q_i(y_i,\theta)\right|^2\right]<\infty \tag{7.15}$$

according to Lemma 7.1, we have, $\forall\theta\in\boldsymbol{\Theta}$,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\left(q_i(y_i,\theta)-\mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right)=0\quad a.s.$$

Let

$$Y_i(y_i,\theta) \triangleq q_i(y_i,\theta) - \mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right] \tag{7.16}$$

On the other hand, $\forall\theta',\theta'' \in \mathbf{\Theta}$, $\theta' < \theta''$ and $\theta^* \in (\theta',\theta'')$, we have

$$\left|\frac{1}{n}\sum_{i=1}^{n} Y_i(y_i,\theta') - \frac{1}{n}\sum_{i=1}^{n} Y_i(y_i,\theta'')\right| \leq |\theta' - \theta''| \left|\frac{\partial}{\partial\theta}\frac{1}{n}\sum_{i=1}^{n} Y_i(y_i,\theta)\right|_{\theta=\theta^*}\Bigg| \tag{7.17}$$

Since $\forall\theta \in \mathbf{\Theta}$

$$\frac{1}{n}\sum_{i=1}^{n} Y_i(y_i,\theta) \xrightarrow{a.s.} 0 \tag{7.18}$$

provided $\sup_i \mathbb{E}[|q_i(y_i,\theta)|^2] < \infty$, therefore

$$\limsup_n \left|\frac{\partial}{\partial\theta}\frac{1}{n}\sum_{i=1}^{n} Y_i(y_i,\theta)\right|_{\theta=\theta^*}\Bigg| < \infty \tag{7.19}$$

holds.

Then by the Theorem 21.10 in [114], $\frac{1}{n}\sum_{i=1}^{n} Y_i(y_i,\theta)$ is S.S.E.

Therefore, in conclusion,

$$\sup_{\theta\in\mathbf{\Theta}} \left|\frac{1}{n}\sum_{i=1}^{n}\left(q_i(y_i,\theta) - \mathbb{E}\left[q_i(y_i,\theta)|\mathcal{F}_{i-1}\right]\right)\right| \xrightarrow{a.s.} 0$$

holds.

□

In order to show that, $\forall\theta \in \mathbf{\Theta}$,

$$\ell(x_i,\theta,\hat{\theta}_{i-1}) - \ell(x_i,\theta_T,\hat{\theta}_{i-1}) - \mathbb{E}_{\theta_T}[\ell(x_i,\theta,\hat{\theta}_{i-1}) - \ell(x_i,\theta_T,\hat{\theta}_{i-1})|\mathcal{F}_{i-1}] \tag{7.20}$$

satisfies the condition of Theorem 7.3, it is sufficient to show that

$$\sup_i \mathbb{E}_{\theta_T}\left[\left|\ell(x_i,\theta,\hat{\theta}_{i-1}) - \ell(x_i,\theta_T,\hat{\theta}_{i-1})\right|^2\right] < \infty \tag{7.21}$$

since

$$\begin{aligned}
&\mathbb{E}_{\theta_T}\left[\left|\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\right|^2\right]\\
=&\mathbb{E}_{\theta_T}\left[\left(\log\frac{f(\hat{\theta}_i-\theta_T)}{f(\hat{\theta}_i-\theta)}+z_i^2-\frac{(x_i-\theta)^2}{f(\hat{\theta}_i-\theta)}\right)^2\right]\\
\triangleq&\mathbb{E}_{\theta_T}\left[(\Phi_i+\Psi_i)^2\right]
\end{aligned}$$

where $\Phi_i=\log\frac{f(\hat{\theta}_i-\theta_T)}{f(\hat{\theta}_i-\theta)}$ and $\Psi_i=z_i^2-\frac{(x_i-\theta)^2}{f(\hat{\theta}_i-\theta)}$.

Recall $x_i\sim\mathcal{N}\left(\theta_T,f(\hat{\theta}_i-\theta_T)\right)$, then for finite θ_T and $f(\hat{\theta}_i-\theta_T)$, we have

$$\mathbb{E}_{\theta_T}[x_i^r]<\infty,\ \ r\in\mathbb{Z}^+ \text{ and } r<\infty$$

Thus, by condition C2 and assumption A1, it is not hard to verify that

$$\sup_i\Phi_i^2<\infty,\ \ \sup_i\mathbb{E}_{\theta_T}\left[\Psi_i^2\right]<\infty,\ \ \sup_i\mathbb{E}_{\theta_T}\left[|\Psi_i|\right]<\infty$$

then

$$\begin{aligned}
&\sup_i\mathbb{E}_{\theta_T}\left[(\Phi_i+\Psi_i)^2\right]\\
\leq&\sup_i\mathbb{E}_{\theta_T}\left[\Phi_i^2+\Psi_i^2+2|\Phi_i|\cdot|\Psi_i|\right]\\
\leq&\sup_i\Phi_i^2+\sup_i\mathbb{E}_{\theta_T}\left[\Psi_i^2\right]+\sup_i|\Phi_i|\cdot\sup_i\mathbb{E}_{\theta_T}\left[|\Psi_i|\right]\\
<&\infty
\end{aligned}$$

From Theorem 7.3, by considering A1 and $\{\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\}$ is differentiable a.s. at each point of $\mathbf{\Theta}$, we have

$$\sup_{\theta\in\mathbf{\Theta}}\left|\frac{1}{n}\sum_{i=1}^n\left(\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})-\mathbb{E}_{\theta_T}\left[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})|\mathcal{F}_{i-1}\right]\right)\right|\xrightarrow{a.s.}0$$

7.2.2 Assumption A3

First, we should show that θ_T is the Identifiably Unique Maximizer(IUM) of

$$\mathbb{E}_{\theta_T}\left[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\middle|\mathcal{F}_{i-1}\right]\quad\forall\theta\in\mathbf{\Theta} \tag{7.22}$$

then θ_T is the IUM. of

$$\frac{1}{n}\sum_{i=1}^{n}\mathbb{E}_{\theta_T}\left[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\middle|\mathcal{F}_{i-1}\right]$$

Therefore, we have

$$\begin{aligned}&\mathbb{E}_{\theta_T}\left[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\middle|\mathcal{F}_{i-1}\right]\\=&\mathbb{E}\left[-\log f\left(\hat{\theta}_{i-1}-\theta\right)-\frac{(x_i-\theta)^2}{f\left(\hat{\theta}_{i-1}-\theta\right)}+\log f\left(\hat{\theta}_{i-1}-\theta_T\right)+\frac{(x_i-\theta_T)^2}{f\left(\hat{\theta}_{i-1}-\theta_T\right)}\middle|\mathcal{F}_{i-1}\right]\end{aligned} \tag{7.23}$$

Taking the first derivative to (7.23), we have

$$\begin{aligned}&\frac{\partial}{\partial\theta}\left(\mathbb{E}_{\theta_T}\left[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\middle|\mathcal{F}_{i-1}\right]\right)\Big|_{\theta=\theta_T}\\=&\frac{2}{f^3\left(\hat{\theta}_{i-1}-\theta\right)}\mathbb{E}_{\theta_T}\left[-(\theta-x_i)f\left(\hat{\theta}_{i-1}-\theta\right)+(\theta-x_i)^2f'\left(\hat{\theta}_{i-1}-\theta\right)-\right.\\&\left.f^2\left(\hat{\theta}_{i-1}-\theta\right)f'\left(\hat{\theta}_{i-1}-\theta\right)\right]\Big|_{\theta=\theta_T}\\=&0\end{aligned}$$

Thus the θ_T is the maximizer of

$$\mathbb{E}_{\theta_T}\left[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})\middle|\mathcal{F}_{i-1}\right]$$

by considering that

$$\begin{aligned}&\frac{\partial^2}{\partial\theta^2}\left(\mathbb{E}_{\theta_T}[\ell(x_i,\theta,\hat{\theta}_{i-1})-\ell(x_i,\theta_T,\hat{\theta}_{i-1})|\mathcal{F}_{i-1}]\right)\Big|_{\theta=\theta_T}\\=&\frac{f\left(\hat{\theta}_{i-1}-\theta_T\right)}{f^3\left(\hat{\theta}_{i-1}-\theta_T\right)}\left(f'^2\left(\hat{\theta}_{i-1}-\theta_T\right)-f\left(\hat{\theta}_{i-1}-\theta_T\right)\left(f''\left(\hat{\theta}_{i-1}-\theta_T\right)+2\right)\right.\end{aligned}$$

$$\mathbb{E}_{\theta_T}\left[((\theta_T - x_i)^2 f''\left(\hat{\theta}_{i-1} - \theta_T\right) + 4(\theta_T - x_i) f'\left(\hat{\theta}_{i-1} - \theta_T\right)) - 2(\theta_T - x_i)^2 \frac{f'^2\left(\hat{\theta}_{i-1} - \theta_T\right)}{f\left(\hat{\theta}_{i-1} - \theta_T\right)}\right]\Bigg)$$
$$= -\frac{2}{f\left(\hat{\theta}_{i-1} - \theta_T\right)} - \left(\frac{f'\left(\hat{\theta}_{i-1} - \theta_T\right)}{f\left(\hat{\theta}_{i-1} - \theta_T\right)}\right)^2$$
$$< 0$$

Let $\theta \in \boldsymbol{\Theta} \setminus \{\theta_T\}$. To prove θ_T is the IUM of (7.22), we have

$$\begin{aligned}
&\mathbb{E}_{\theta_T}\left[\ell(x_i, \theta_T, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1}\right] - \mathbb{E}_{\theta_T}\left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1}\right] \\
=&0 - \mathbb{E}\left[-\log f\left(\hat{\theta}_{i-1} - \theta\right) - \frac{(x_i - \theta)^2}{f\left(\hat{\theta}_{i-1} - \theta\right)} + \log f\left(\hat{\theta}_{i-1} - \theta_T\right) + \frac{(x_i - \theta_T)^2}{f\left(\hat{\theta}_{i-1} - \theta_T\right)} \middle| \mathcal{F}_{i-1}\right] \\
=&\frac{\theta^2 + f\left(\hat{\theta}_{i-1} - \theta_T\right)}{f\left(\hat{\theta}_{i-1} - \theta\right)} - \log \frac{f\left(\hat{\theta}_{i-1} - \theta_T\right)}{f\left(\hat{\theta}_{i-1} - \theta\right)} - 1 \\
>&\frac{f^2\left(\hat{\theta}_{i-1} - \theta_T\right)}{f^2\left(\hat{\theta}_{i-1} - \theta\right)} - \log \frac{f^2\left(\hat{\theta}_{i-1} - \theta_T\right)}{f^2\left(\hat{\theta}_{i-1} - \theta\right)} - 1
\end{aligned}$$

Since $f^2\left(\|\hat{\theta}_{i-1} - \theta\|, \sigma^2\right) \neq f^2\left(\|\hat{\theta}_{i-1} - \theta_T\|, \sigma^2\right)$, thus

$$\frac{f^2\left(\hat{\theta}_{i-1} - \theta\right)}{f^2\left(\hat{\theta}_{i-1} - \theta_T\right)} - \log \frac{f^2\left(\hat{\theta} - \theta\right)}{f^2\left(\hat{\theta} - \theta_T\right)} - 1 > 0$$

Thus, θ_T is the IUM of (7.22) and also the IUM of

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{E}_{\theta_T}\left[\ell(x_i, \theta, \hat{\theta}_{i-1}) - \ell(x_i, \theta_T, \hat{\theta}_{i-1}) \middle| \mathcal{F}_{i-1}\right]$$

Therefore, A1, A2 and A3 in Theorem 7.1 are satisfied, as a conclusion, the MLE of robotic problem with Gaussian noise is consistent provided the conditions C1 – C3 listed in Section 6.2 on the variance are satisfied.

7.3 Conclusions

In this chapter, we investigate the consistency of MLE in a complicated situation that the mean and variance of measurement are dependent on the unknown parameter and more-

over, the estimate relies on the all previous information. This situation is of interest in the engineering application and theoretical analysis. The key theorem is given in Theorem 7.1 which is the martingale version of general consistency theorem. In Section 7.2.1 and 7.2.2, the likelihood function of underlying problem is shown to satisfy the assumptions A1, A2 and A3 listed in Theorem 7.1, and therefore, the MLE is consistent.

Part IV

Conclusion and Future Work

Two critical issues, the phase ambiguity and finding a position using distance dependent noise, in localization are considered in this thesis in Part II and III respectively.

In Part II, the improved efficient CRT and lattice algorithms in resolving ambiguity are presented to fit more practical situations, such as providing more flexibility to the designers and improving the performance of the conventional algorithm in terms of increasing reconstruction probability and decreasing computational complexity. On the other hand, in Part III, the consistency problem of target localization with distance dependent noise is discussed. Two cases are considered and the consistency results are given under some conditions. The author of this thesis is keen to continue his research along these topics. In particular, he has great interest to address the following open issues with effort towards the following open questions towards designing efficient algorithm and theoretical analysis.

1. The CRT-related algorithms in addressing the noisy phase ambiguity problem require that the used wavelengths satisfy some conditions, e.g. the coprimality condition. However, in practice, these conditions cannot always be satisfied due to the limitation of available wavelengths range. This could prevent the applications of CRT in real world. Therefore, efficient algorithms using arbitrary wavelengths are preferred and the performance analysis is required.

2. An efficient lattice algorithm using arbitrary wavelengths is presented in this Section 4.3 and its efficiency is demonstrated via simulation in both reconstruction probability and computational complexity. However, the theoretical analysis of the performance is pivotal in applying this algorithm. Therefore, in the future work, the reconstruction probability and computational complexity of lattice algorithm using arbitrary wavelengths should be analyzed.

3. The consistency result of robotic problem with distance dependent noise when robot is movable is given in Chapter 7. In the proof, a crucial condition is that the parameter space of unknown target's location is compact. However, in theory, it is interesting to prove that the MLE will lie in a compact set eventually without any assumption on the parameter space. This problem is found to be very challeng-

ing even in 1 dimension case. Therefore, in the future, one of the main job is to find a proper method to solve this problem.

4. The behaviors of MLE of the robotic problem are discussed in Chapter 6 and 7. There are some other interesting questions about Bayesian inference. The first problem is to find a conjugate prior for updating the posterior density and, the second one is to prove the consistency for the Bayesian inference using maximizing posterior density or posterior mean.

Bibliography

[1] T. Ahmad, X. J. Li, and B.-C. Seet, "Parametric loop division for 3d localization in wireless sensor networks," *Sensors*, vol. 17, no. 7, 2017, p. 1697.

[2] N. Iliev and I. Paprotny, "Review and comparison of spatial localization methods for low-power wireless sensor networks," *IEEE Sensors Journal*, vol. 15, no. 10, Oct. 2015, pp. 5971–5987.

[3] J. Kuriakose, S. Joshi, R. V. Raju, and A. Kilaru, "A review on localization in wireless sensor networks," in *Advances in Signal Processing and Intelligent Recognition Systems*, A. G. S. Thampi and J. Mukhopadhyay, Eds. Switzerland: Springer, Cham, 2014, vol. 264, pp. 599–610.

[4] C. Hekimian-Williams, B. Grant, X. Liu, Z. Zhang, and P. Kumar, "Accurate localization of rfid tags using phase difference," in *2010 IEEE International Conference on RFID (IEEE RFID 2010)*, Orlando, FL, USA, Apr. 2010, pp. 89–96.

[5] S. Gezici, Z. Tian, G. B. Giannakis, H. Kobayashi, A. F. Molisch, H. V. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, Jul. 2005, pp. 70–84.

[6] S. Lanzisera, D. Zats, and K. S. J. Pister, "Radio frequency time-of-flight distance measurement for low-cost wireless sensor localization," *IEEE Sensors Journal*, vol. 11, no. 3, Mar. 2011, pp. 837–845.

[7] L. Cheng, C. Wu, Y. Zhang, H. Wu, M. Li, and C. Maple, "A survey of localization in wireless sensor network," *International Journal of Distributed Sensor Networks*, vol. 8, no. 12, 2012.

[8] N. Bulusu, J. Heidemann, D. Estrin, and T. Tran, "Self-configuring localization systems: Design and experimental evaluation," *ACM Trans. Embed. Comput. Syst.*, vol. 3, no. 1, Feb. 2004, pp. 24–60.

[9] C. Liu, K. Wu, and T. He, "Sensor localization with ring overlapping based on comparison of received signal strength indicator," in *2004 IEEE International Conference on Mobile Ad-hoc and Sensor Systems (IEEE Cat. No.04EX975)*, Oct. 2004, pp. 516–518.

[10] T. He, C. Huang, B. M. Blum, J. A. Stankovic, and T. Abdelzaher, "Range-free localization schemes for large scale sensor networks," in *Proceedings of the 9th Annual International Conference on Mobile Computing and Networking*, ser. MobiCom '03. New York, NY, USA: ACM, 2003, pp. 81–95.

[11] K. Langendoen and N. Reijers, "Distributed localization in wireless sensor networks: A quantitative comparison," *Computer Networks*, vol. 43, no. 4, Nov. 2003, pp. 499–518.

[12] G. Mao, B. Fidan, and B. D. Anderson, "Wireless sensor network localization techniques," *Computer Networks*, vol. 51, no. 10, 2007, pp. 2529 – 2553.

[13] M. Maróti, P. Völgyesi, S. Dóra, B. Kusý, A. Nádas, A. Lédeczi, G. Balogh, and K. Molnár, "Radio interferometric geolocation," in *Proceedings of the 3rd International Conference on Embedded Networked Sensor Systems*, ser. SenSys '05. New York, NY, USA: ACM, 2005, pp. 1–12.

[14] A. Loeffler, "Localizing passive UHF RFID tags with wideband signals," in *2011 IEEE International Conference on Microwaves, Communications, Antennas and Electronic Systems (COMCAS 2011)*, Nov. 2011, pp. 1–6.

[15] Y. Cheng, X. Wang, T. Caelli, and B. Moran, "Tracking and localizing moving targets in the presence of phase measurement ambiguities," *IEEE Transactions on Signal Processing*, vol. 59, no. 8, Aug. 2011, pp. 3514–3525.

[16] Z. Huang, G. Zhan, D. Ying, R. Zhou, J. Pan, and Y. Yan, "Robust multiple speech source localization based on phase difference regression," in *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Oct. 2016, pp. 1–5.

[17] M. Ilyas and I. Mahgoub, *Handbook of Sensor Networks: Compact Wireless and Wired Sensing Systems*. CRC Press, 2004.

[18] K. Römer, "Time synchronization and localization in sensor networks," Ph.D. dissertation, Swiss Federal Institute of Technology Zurich, 2005.

[19] A. Makki, A. Siddig, M. Saad, J. R. Cavallaro, and C. J. Bleakley, "Indoor localization using 802.11 time differences of arrival," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, Mar. 2016, pp. 614–623.

[20] D. Da-Nan and B. Yehuda, "Global positioning system network analysis with phase ambiguity resolution applied to crustal deformation studies in california," *Journal of Geophysical Research: Solid Earth*, vol. 94, no. B4, 1989, pp. 3949–3966.

[21] R. Shimoyama and K. Yamazaki, "Multiple acoustic source localization using ambiguous phase differences under reverberative conditions," *Acoustical Science and Technology*, vol. 25, no. 6, 2004, pp. 446–456.

[22] A. Ledeczi, P. Volgyesi, J. Sallai, B. Kusy, X. Koutsoukos, and M. Maroti, "Towards precise indoor rf localization," in *5th Workshop Embedded Networked Sensors (HotEmNets)*, Charlottesville, Virginia, USA, Nov. 2008.

[23] G. Mao, *Localization Algorithms and Strategies for Wireless Sensor Networks: Monitoring and Surveillance Techniques for Target Tracking*, ser. Premier reference source. Information Science Reference, 2009.

[24] W. Wang and X.-G. Xia, "A closed-form robust chinese remainder theorem and its performance analysis," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, Nov. 2010, pp. 5655–5666.

[25] W. Li, X. Wang, and B. Moran, "Wireless signal travel distance estimation using non-coprime wavelengths," *IEEE Signal Processing Letters*, vol. 24, no. 1, Jan. 2017, pp. 27–31.

[26] ——, "Efficient distance estimation using wrapped phase measurements," in *IET International Radar Conference 2015*, Hangzhou, China, Oct. 2015, pp. 1–6.

[27] W. Li, X. Wang, X. Wang, and B. Moran, "Distance estimation using wrapped phase measurements in noise," *IEEE Transactions on Signal Processing*, vol. 61, no. 7, Apr. 2013, pp. 1676–1688.

[28] D. Vasisht, S. Kumar, and D. Katabi, "Decimeter-level localization with a single wifi access point," in *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, Santa Clara, CA, Mar. 2016.

[29] C. M. Alabaster, E. J. Hughes, and J. H. Matthew, "Medium prf radar prf selection using evolutionary algorithms," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 3, Jul. 2003, pp. 990–1001.

[30] W. Li, X. Wang, and B. Moran, "A lattice algorithm for optimal phase unwrapping in noise," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2016)*, Shanghai, China, Mar. 2016, pp. 2896–2900.

[31] R. G. McKilliam, A. Grant, and I. V. L. Clarkson, "Finding a closest point in a lattice of voronoi's first kind," *SIAM Journal on Discrete Mathematics*, vol. 28, no. 3, 2014, pp. 1405–1422.

[32] W. Li, X. Wang, and B. Moran, "A lattice method for resolving range ambiguity in dual-frequency rfid tag localisation," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2017)*, New Orleans, USA, Mar. 2017, pp. 3156–3160.

[33] ——, "An improved chinese remainder theorem for range estimation," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP2018)*, Accepted 2018.

[34] ——, "An improved lattice algorithm and its performance analysis," will submit to IEEE Signal Processing Letter.

[35] K. Aardal and F. Eisenbrand, "Integer programming, lattices, and results in fixed dimension," in *Discrete Optimization*, ser. Handbooks in Operations Research and Management Science, K. Aardal, G. Nemhauser, and R. Weismantel, Eds. Elsevier, 2005, vol. 12, no. Supplement C, pp. 171 – 243.

[36] W. Schmidt, *Diophantine Approximation*, ser. Bonner Meteorologische Abhandlungen. Springer, 1996.

[37] A. Schrijver, *Theory of Linear and Integer Programming*. New York, NY, USA: John Wiley & Sons, Inc., 1986.

[38] O. Regev, "Lattice-based cryptography," in *Proceedings of the 26th Annual International Conference on Advances in Cryptology*, ser. CRYPTO'06. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 131–141.

[39] J. Hoffstein, J. Pipher, and J. Silverman, *An Introduction to Mathematical Cryptography*, ser. Undergraduate Texts in Mathematics. Springer New York, 2014.

[40] R. Zamir, *Lattice Coding for Signals and Networks: A Structured Coding Approach to Quantization, Modulation, and Multiuser Information Theory*. New York, NY, USA: Cambridge University Press, 2014.

[41] U. Erez, S. Litsyn, and R. Zamir, "Lattices which are good for (almost) everything," *IEEE Transactions on Information Theory*, vol. 51, no. 10, Oct. 2005, pp. 3401–3416.

[42] N. Sommer, M. Feder, and O. Shalvi, "Low-density lattice codes," *IEEE Transactions on Information Theory*, vol. 54, no. 4, Apr. 2008, pp. 1561–1585.

[43] N. Benvenuto and G. Cherubini, *Algorithms for Communications Systems and Their Applications*. Wiley, 2002.

[44] R. G. McKilliam, B. G. Quinn, I. V. L. Clarkson, and B. Moran, "Frequency estimation by phase unwrapping," *IEEE Transactions on Signal Processing*, vol. 58, no. 6, Jun. 2010, pp. 2953–2963.

[45] H. W. Lenstra, "Integer programming with a fixed number of variables," *Mathematics of Operations Research*, vol. 8, no. 4, 1983, pp. 538–548.

[46] H. W. Lenstra, A. K. Lenstra, and L. Lovász, "Factoring polynomials with rational coefficients." *Mathematische Annalen*, vol. 261, 1982, pp. 515–534.

[47] A. M. Odlyzko, "The rise and fall of knapsack cryptosystems," in *In Cryptology and Computational Number Theory*. A.M.S, 1990, pp. 75–88.

[48] J. Antoine and J. Stern, "Lattice reduction: A toolbox for the cryptanalyst," *Journal of Cryptology*, vol. 11, no. 3, 1998, pp. 161–185.

[49] P. Q. Nguyen and J. Stern, "The two faces of lattices in cryptology," in *Revised Papers from the International Conference on Cryptography and Lattices*, ser. CaLC '01. London, UK, UK: Springer-Verlag, 2001, pp. 146–180.

[50] H. Minkowski, "Ueber die positiven quadratischen formen und über kettenbruchähnliche algorithmen." *Journal für die reine und angewandte Mathematik*, vol. 107, 1891, pp. 278–297.

[51] A. Korkine and G. Zolotareff, "Sur les formes quadratiques," *Mathematische Annalen*, vol. 6, no. 3, 1873, pp. 366–389.

[52] J. Cassels, *An Introduction to the Geometry of Numbers*, ser. Classics in Mathematics. Springer Berlin Heidelberg, 1996.

[53] P. Q. Nguyen and D. Stehlé, "Low-dimensional lattice basis reduction revisited," *ACM Trans. Algorithms*, vol. 5, no. 4, Nov. 2009, pp. 46:1–46:48.

[54] B. L. van der Waerden, "Die reduktionstheorie der positiven quadratischen formen," *Acta Mathematica*, no. 96, 1968, pp. 265–309.

[55] A. Korkine and G. Zolotareff, "Sur les formes quadratiques positives quaternaires," *Mathematische Annalen*, vol. 5, no. 4, 1872, pp. 581–583.

[56] P. Q. Nguyen and B. Valle, *The LLL Algorithm: Survey and Applications*, 1st ed. Springer Publishing Company, Incorporated, 2009.

[57] M. R. Bremner, *Lattice Basis Reduction: An Introduction to the LLL Algorithm and Its Applications*, 1st ed. Boca Raton, FL, USA: CRC Press, Inc., 2011.

[58] E. Viterbo and E. Biglieri, "Computing the Voronoi cell of a lattice: the diamond-cutting algorithm," *IEEE Transactions on Information Theory*, vol. 42, no. 1, Jan. 1996, pp. 161–171.

[59] L. Babai, "On lovász' lattice reduction and the nearest lattice point problem," *Combinatorica*, vol. 6, no. 1, 1986, pp. 1–13.

[60] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Transactions on Information Theory*, vol. 48, no. 8, Aug. 2002, pp. 2201–2214.

[61] B. Hassibi and H. Vikalo, "On the sphere-decoding algorithm i. expected complexity," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, Aug. 2005, pp. 2806–2818.

[62] H. Vikalo and B. Hassibi, "On the sphere-decoding algorithm ii. generalizations, second-order statistics, and applications to communications," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, Aug. 2005, pp. 2819–2834.

[63] N. Sommer, M. Feder, and O. Shalvi, "Finding the closest lattice point by iterative slicing," *SIAM Journal on Discrete Mathematics*, vol. 23, no. 2, 2009, pp. 715–731.

[64] B. Kusy, A. Ledeczi, M. Maroti, and L. Meertens, "Node-density independent localization," in *2006 5th International Conference on Information Processing in Sensor Networks*, April 2006, pp. 441–448.

[65] B. Kusy, J. Sallai, G. Balogh, A. Ledeczi, V. Protopopescu, J. Tolliver, F. DeNap, and M. Parang, "Radio interferometric tracking of mobile wireless nodes," in *Proceedings of the 5th International Conference on Mobile Systems, Applications and Services*, ser. MobiSys '07. New York, NY, USA: ACM, 2007, pp. 139–151.

[66] Y. Wang, L. Li, X. Ma, M. Shinotsuka, C. Chen, and X. Guan, "Dual-tone radio interferometric positioning systems using undersampling techniques," *IEEE Signal Processing Letters*, vol. 21, no. 11, Nov. 2014, pp. 1311–1315.

[67] Y. Wang, X. Ma, C. Chen, and X. Guan, "Designing dual-tone radio interferometric positioning systems," *IEEE Transactions on Signal Processing*, vol. 63, no. 6, Mar. 2015, pp. 1351–1365.

[68] M. Shinotsuka, Y. Wang, X. Ma, and G. T. Zhou, "Analyzing radio interferometric positioning systems with undersampling receivers," *EURASIP Journal on Advances in Signal Processing*, vol. 2015, no. 1, Sep. 2015, p. 82.

[69] X. Wang, B. Moran, and M. Brazil, "Hyperbolic positioning using rips measurements for wireless sensor networks," in *2007 15th IEEE International Conference on Networks*, Nov 2007, pp. 425–430.

[70] C. Ding, D. Pei, and A. Salomaa, *Chinese Remainder Theorem: Applications in Computing, Coding, Cryptography*. River Edge, NJ, USA: World Scientific Publishing Co., Inc., 1996.

[71] L. N.Childs, "The chinese remainder theorem," in *A Concrete Introduction to Higher Algebra. Undergraduate Texts in Mathematics*, ser. Undergraduate Texts in Mathematics, L. N.Childs, Ed. Springer, New York, NY, 1995.

[72] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, 2nd ed. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc., 1994.

[73] X.-G. Xia, "An efficient frequency-determination algorithm from multiple undersampled waveforms," *IEEE Signal Processing Letters*, vol. 7, no. 2, Feb. 2000, pp. 34–37.

[74] X. Li, H. Liang, and X.-G. Xia, "A robust chinese remainder theorem with its applications in frequency estimation from undersampled waveforms," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, Nov. 2009, pp. 4314–4322.

[75] W. Wang, X. Li, W. Wang, and X.-G. Xia, "Maximum likelihood estimation based robust chinese remainder theorem for real numbers and its fast algorithm," *IEEE Transactions on Signal Processing*, vol. 63, no. 13, Jul. 2015, pp. 3317–3331.

[76] L. Xiao, X.-G. Xia, and W. Wang, "Multi-stage robust chinese remainder theorem," *IEEE Transactions on Signal Processing*, vol. 62, no. 18, Sept. 2014, pp. 4772–4785.

[77] L. Xiao, X.-G. Xia, and H. Huo, "New conditions on achieving the maximal possible dynamic range for a generalized chinese remainder theorem of multiple integers," *IEEE Signal Processing Letters*, vol. 22, no. 12, Dec. 2015, pp. 2199–2203.

[78] X. Li, X.-G. Xia, W. Wang, and W. Wang, "A robust generalized chinese remainder theorem for two integers," *IEEE Transactions on Information Theory*, vol. 62, no. 12, Dec. 2016, pp. 7491–7504.

[79] Baldoni, M. Welleda, C. Ciliberto, and G. M. Piacentini Cattaneo, *Elementary number theory, Cryptography and Codes*. Berlin, Heidelberg: Springer Verlag, 2009.

[80] H. Cohen, *A Course in Computational Algebraic Number Theory*. New York, NY, USA: Springer-Verlag New York, Inc., 1993.

[81] X. Li, H. Liang, and X.-G. Xia, "A robust chinese remainder theorem with its applications in frequency estimation from undersampled waveforms," *IEEE Transactions on Signal Processing*, vol. 57, no. 11, Nov. 2009, pp. 4314–4322.

[82] X. Li, Y. Zhang, and M. G. Amin, "Multifrequency-based range estimation of rfid tags," in *IEEE International Conference on Rfid*, 2009, pp. 147–154.

[83] F. Beichelt, *Stochastic Processes in Science, Engineering and Finance*. CRC Press, 2006.

[84] K. Chung and Z. Zhao, *From Brownian Motion to Schrödingers Equation*, ser. Grundlehren der mathematischen Wissenschaften. Springer Berlin Heidelberg, 2001.

[85] R. Bass, *Probabilistic Techniques in Analysis*, ser. Probability and Its Applications. Springer New York, 1994.

[86] D. Duffie, *Dynamic Asset Pricing Theory*. Princeton University Press, 1996.

[87] G. N. Saridis, "Comparison of six on-line identification algorithms," *Automatica*, vol. 10, no. 1, 1974, pp. 69 – 79.

[88] G. Saridis, *Self-organizing control of stochastic systems*, ser. Control and systems theory. M. Dekker, 1977.

[89] A. Segall, *Martingale Approach to Modeling, Estimation and Detection of Jump Processes*. Stanford University, 1973.

[90] J. Doob, *Stochastic Processes*, ser. Wiley Publications in Statistics. John Wiley & Sons, 1953.

[91] I. V. Basawa, P. D. Feigin, and C. C. Heyde, "Asymptotic properties of maximum likelihood estimators for stochastic processes," *Sankhy: The Indian Journal of Statistics, Series A (1961-2002)*, vol. 38, no. 3, 1976, pp. 259–270.

[92] P. Hall, E. Lukacs, Z. W. Birnbaum, and C. C. Heyde, "Martingale limit theory and its application," *Journal of the American Statistical Association*, vol. 78, no. 78, 1980, pp. 301–308.

[93] P. Hall and C. C. Heyde, *Martingale limit theory and its application*. Academic press, 2014.

[94] D. Williams, *Probability with Martingales*. Cambridge university press, 1991.

[95] A. Gut, *Probability: a graduate course*. Springer Science & Business Media, 2012, vol. 75.

[96] X. Wang, M. Fu, and H. Zhang, "Target tracking in wireless sensor networks based on the combination of kf and mle using distance measurements," *IEEE Transactions on Mobile Computing*, vol. 11, no. 4, 2012, pp. 567–576.

[97] N. Patwari, J. N. Ash, S. Kyperountas, I. Hero, Alfred O., R. L. Moses, and N. S. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Magazine*, vol. 22, no. 4, 2005, pp. 54–69.

[98] A. De, A. Ribeiro, W. Moran, and D. E. Koditschek, "Convergence of bayesian histogram filters for location estimation," in *Decision and Control*, 2014, pp. 7047 – 7053.

[99] J. L. Doob, "Probability and statistics," *Transactions of the American Mathematical Society*, vol. 36, no. 4, 1934, pp. 759–775.

[100] H. Cramer, "Mathematical methods of statistics," *Princeton Mathematical*, vol. 53, no. 4, 1946, pp. 431–450.

[101] A. Wald, "Asymptotic properties of the maximum likelihood estimate of an unknown parameter of a discrete stochastic process," *Annals of Mathematical Statistics*, vol. 19, no. 1, 1948, pp. 40–46.

[102] Y. Bar-Shalom, "On the asymptotic properties of the maximum-likelihood estimate obtained from dependent observations," *Journal of the Royal Statistical Society*, vol. 33, no. 1, 1971, pp. 72–77.

[103] B. R. Bhat, "On the method of maximum-likelihood for dependent observations," *Journal of the Royal Statistical Society*, vol. 36, no. 1, 1974, pp. 48–53.

[104] M. J. Crowder, "Maximum likelihood estimation for dependent observations," *Journal of the Royal Statistical Society*, vol. 38, no. 1, 1976, pp. 45–53.

[105] K. McGoff, S. Mukherjee, A. Nobel, N. Pillai *et al.*, "Consistency of maximum likelihood estimation for some dynamical systems," *The Annals of Statistics*, vol. 43, no. 1, 2015, pp. 1–29.

[106] R. D. H. Heijmans and J. Magnus, "Asymptotic normality of maximum likelihood estimators obtained from normally distributed but dependent observations," *Econometric Theory*, vol. 2, no. 3, 1986, pp. 374–412.

[107] T. Bollerslev and J. M. Wooldridge, "Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances," *Econometric Reviews*, vol. 11, no. 2, 1992, pp. 143–172.

[108] T. Ferguson, *A Course in Large Sample Theory*, ser. Chapman & Hall Texts in Statistical Science Series. Taylor & Francis, 1996.

[109] R. Silverman, *Essential Calculus with Applications*, ser. Contemporary mathematics. Dover Publications, 1989.

[110] W. K. Newey and D. McFadden, "Large sample estimation and hypothesis testing," in *Handbook of Econometrics*, R. Engle and D. McFadden, Eds. Amsterdam, The Netherlands: Elsevier, 1986, vol. IV, ch. 36, pp. 2111–2245.

[111] A. van der vaart and J. Wellner, *Weak Convergence and Empirical Processes: With Applications to Statistics*, ser. Springer Series in Statistics. Springer New York, 2013.

[112] Y. S. Chow, "On a strong law of large numbers for martingales," *The Annals of Mathematical Statistics*, vol. 38, no. 2, Apr. 1967, p. 610.

[113] D. W. K. Andrews, "Generic uniform convergence," *Econometric Theory*, vol. 8, no. 2, 1992, pp. 241–257.

[114] J. Davidson, *Stochastic Limit Theory: An Introduction for Econometricians*, ser. Advanced Texts in Econometrics. OUP Oxford, 1994.