Localization and Communication
for UWB-based Wireless Sensor Networks

Yiyin Wang

Localization and Communication
for UWB-based Wireless Sensor Networks

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof.ir. K.C.A.M. Luyben,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op maandag 14 november 2011 om 12:30 uur
door

Yiyin WANG
Master of Science in Microelectronics
geboren te Chengdu, China.

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. A.-J. van der Veen

Copromotor:
Dr. ir. G.J.T. Leus

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. A.-J. van de Veen	Technische Universiteit Delft, promotor
Dr. ir. G.J.T. Leus	Technische Universiteit Delft, copromotor
Prof. dr. K.G. Langendoen	Technische Universiteit Delft
Prof. dr. O. Yarovyi	Technische Universiteit Delft
Prof. dr. P.J.M. Havinga	Universiteit Twente
Prof. dr. Xiaoli Ma	Georgia Institute of Technology
Dr. ir. Z. Papp	TNO

In memory of my father, Shipei Wang

and to my mother, Lianfang

my parents-in-law, Zhenxiang and Songwei

and my husband, Xi

Summary

The great demand for location-aware wireless sensor networks (WSNs) motivates the research in this thesis. The unique characteristics of WSNs impose numerous challenges on localization and communication. In this thesis, we handle some key challenges and provide affordable solutions.

Impulse radio ultra wideband (IR-UWB) is employed as the fundamental technology for both localization and communication due to its distinctive advantages in accurate ranging and reliable communication. The following aspects are treated in this thesis.

- *Transmitted-reference (TR) UWB communication systems:* IR-UWB processing in the digital domain usually asks for very high sampling rates. The TR-UWB scheme allows for sub-Nyquist rate sampling by correlating the received pulse sequence with its delayed version in the analog domain. Thus, it avoids the daunting Nyquist sampling rate, relaxes the stringent synchronization requirements, and only asks for aggregate channel coefficients. A data model including various kinds of interferences is employed, and then a complete receiver is proposed including signal detection, channel estimation, synchronization and equalization.

- *Theoretical ranging bounds and practical ranging methods based on IR-UWB:* We investigate the theoretical ranging accuracy of a novel method, which exploits the range information in both the amplitude and the time delay of the received signal. The investigations are conducted not only for an additive white Gaussian noise (AWGN) channel with attenuation, but also for an AWGN channel with both attenuation and shadowing. Furthermore, a practical ranging method based on time-of-arrival (TOA) estimation using UWB IRs is developed. Stroboscopic sampling is employed to sacrifice transmission efficiency for a lower sampling rate. Moreover, it can maintain the same ranging resolution as Nyquist sampling can achieve. Due to the long preamble required by stroboscopic sampling, the clock drift, which is an accumulative effect over time caused by the relative clock skew between

different clocks, is one of the main error sources in TOA estimation. Therefore, TOA estimation methods with clock drift calibration are explored to dramatically mitigate the influence of the drift.

- *Various localization and tracking methods:*

  - *Extended multi-dimensional scaling (MDS):* Since the classical MDS cannot be applied to general networks with missing links, we extend the classical MDS algorithm to deal with a special kind of network with specific missing links. Our goal is to jointly estimate the positions of all the nodes given partial pairwise distance measurements up to a translation, rotation, and reflection.

  - *Reference-free time-based localization:* Low-complexity least-squares (LS) estimators based on time-of-arrival (TOA) or time-difference-of-arrival (TDOA) measurements have been developed in literature to locate a target node with the help of anchors (nodes with known positions). They require to select a reference anchor in order to cancel nuisance parameters or relax stringent synchronization requirements, and suffer from a poor reference selection. We propose reference-free localization estimators based on TOA measurements to decouple the reference dependency. Furthermore, we generalize existing reference-based closed-form localization estimators using TOA or TDOA measurements, and shed new light on their relations to clarify some confusions that still persist in recent literature.

  - *Robust time-based localization:* Time-based localization approaches attract a lot of interest due to their high accuracy and potentially low cost for WSNs. However, time-based localization is tightly coupled with clock synchronization. Thus, the reliability of timestamps in time-based localization becomes an important yet challenging task to deal with. Regardless of the reliability of the timestamps from the target node, we propose a novel ranging protocol, namely asymmetric trip ranging (ATR), which leads to localization methods that are naturally immune to internal attacks mounted by a compromised target node. Robust localization strategies using the ATR protocol based on TOA measurements are proposed to localize a target node with the help of anchors for asynchronous networks.

  - *Kalman tracking:* Due to the nonlinearity of the localization problem, a Kalman filter (KF) is usually replaced by an extended KF (EKF) for tracking a mobile target. However, the modeling errors inherently contained in the EKF degrade the tracking performance. Therefore, we make use of the ATR protocol again, carry out exact linearizations, and achieve a KF based on a linear measurement model to track a mobile target with the aid of fixed anchors.

Contents

Introduction

1.1 Location-aware wireless sensor networks

Earthquake in Japan: at 14:46 local time March 11th, 2011, the north-east coast of Japan was hit by the most powerful earthquake since records, which triggered a massive tsunami as a result. Thousands of people were dead, and many more people were missing. Around 60 international teams from 45 countries have been placed on alert to assist Japanese teams to rescue victims trapped beneath collapsed buildings.

Jewelry thieves in Tefaf: in March 20th, 2011, jewelry thieves struck again during the Tefaf antiques fair, which is the world's leading fine art fair held in Maastricht, The Netherlands. It was not the first time that thieves struck at the fair. In 2010, a pendant and ring with combined value of $860,000$ euros were stolen. In 2008, a necklace valued at 1.2 million euros was stolen.

E. coli outbreak across Europe: in May 2011, the outbreak of a virulent strain of E. coli infected more than 2,200 people in at least 12 countries across Europe. Forty-two victims died in Germany, and one in Sweden according to the report from the European Centre for Disease Prevention and Control in June 23rd 2011. Germany reported that cucumbers from Spain were the suspicious source for the E. coli outbreak. Tens of thousands of kilos of fresh fruit and vegetables grown in Spain were destroyed. Spanish fruit and vegetable exporters estimated a loss of 200 million dollars a week. But later, it was found that sprouts grown in Germany were the likely source. Only until June 30th 2011 the German Federal Institute for Risk Assess-
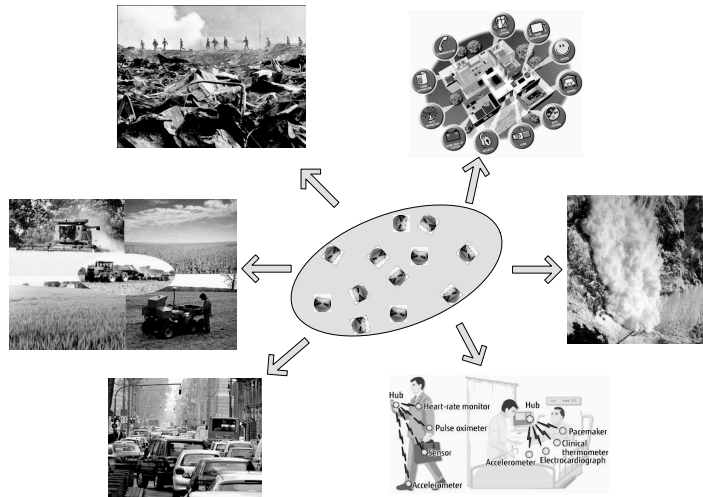
Figure 1.1: Applications of WSNs.

> ment announced that seeds of fenugreek from Egypt were likely the cause
> of the E. coli outbreak.

All these recent events have significantly affected people's lives. We engineers ask
ourselves what we can do to make things better. Could we search for survivors
faster and smarter? Could we prevent the loss of assets? Could we have a better
monitoring and control of our food?

Location-aware wireless sensor networks (WSNs) may offer a promising solu-
tion to all these problems. Exciting advances in IC design have made it possible to
miniaturize sensor devices with measuring, computation, and communication cap-
abilities. Although each sensor has limited capabilities, the deployment of a large
wireless sensor network (WSN) can accomplish numerous complicated functions
in a wide range of applications, such as monitoring and control in environmental,
agricultural or industrial systems, as shown in Fig. 1.1. For instance, WSNs can
provide valuable information for search and rescue during severe disasters, such as
earthquakes and avalanches; asset tracking can be accomplished by WSNs; preci-
sion agriculture can employ WSNs to reduce costs and environmental impact by
fertilizing and watering only where necessary; WSNs can be used to detect pollu-
tion and identify the sources; product quality control in industrial processes can
be achieved by making use of wireless sensors to measure, monitor and actuate;
and traffic systems can take benefit from WSNs to monitor traffic throughput and
direct traffic flows.

The data of the sensors has to be associated with the locations where the data
was measured, and at the same time be accessible via flexible wireless communica-
tions. Therefore, localization and communication have equal primary importance

in WSNs. Furthermore, localization itself is a challenging research topic under investigation for many decades. It finds applications not only in WSNs [1,2], but also in the global positioning system (GPS) [3], radar systems [4], underwater systems [5], acoustic systems [6,7], cellular networks [8], as well as wireless local area networks (WLANs) [9], etc. It is embraced everywhere at any scale. New applications of localization are continuously emerging, which motivates further exploration and attracts many researchers from different research areas, such as geophysics, signal processing, aerospace engineering and computer science. To name a few localization applications, there are geographic routing, warehousing, logistics, etc.

Considering localization and communication for WSNs at the physical layer, we are confronted with a number of unique challenges:

- **Accurate localization and reliable communication in harsh environments**

  In general, WSNs will be deployed in harsh environments, such as a disaster scene, inside process equipment, office buildings and green houses, etc. These environments are normally GPS-denied and rich in multipath channels, which makes it very difficult to obtain accurate location estimates and perform reliable communications. Moreover, the requirements for localization accuracy also vary with the environment as shown in Fig. 1.2. For outdoor environments, meter-level accuracy may satisfy the requirements. On the other hand, for indoor environments, centimeter accuracy is sometimes required.

- **Low cost and low complexity**

  The number of sensors deployed in a WSN can be huge. Therefore, there are stringent constraints upon their cost and complexity. When we develop the localization and communication system for WSNs, we should always bear these constraints in mind and balance them with the obtained performance.

- **Robustness**

  Due to the limited cost, lots of errors are caused by uncertainties of sensor components. For example, sensors may employ poor crystals for clocking, which will introduce severe synchronization problems in both localization and communication. Meanwhile, WSNs are vulnerable to malicious attacks. Robustness has to be taken into account in the system development.

1.2 UWB technology

1.2.1 Overview of UWB

UWB technology is not new for wireless communications. Its history starts from the first impulse radio (IR) experiment conducted by Marconi for wireless tele-
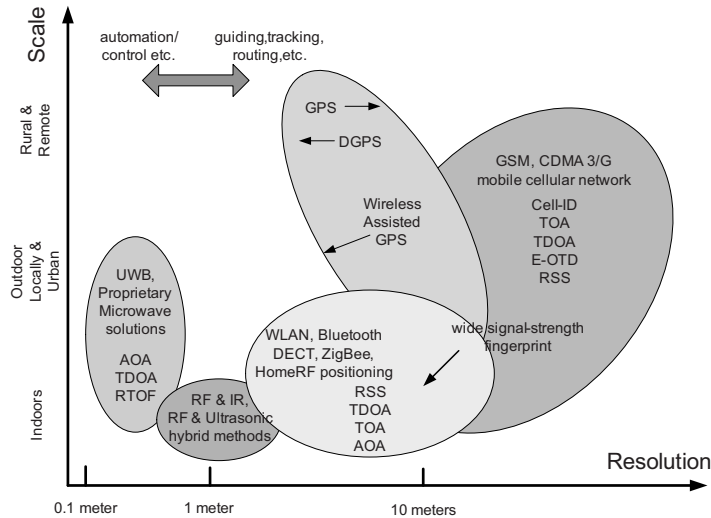
Figure 1.2: Outline of current wireless localization systems [10].

graphy in the late 1890s. Very narrow pulses in the time domain, which occupy a large bandwidth in the frequency domain, are employed. However, the implementation difficulties of IR have pushed wireless communications to an opposite direction, which now employs narrowband (NB) signals to make feasible hardware systems. Driven by the great demand for wireless communications with multiple access and high data rates, code division multiple access (CDMA) and orthogonal frequency-division multiplexing (OFDM) technologies for NB signals are successfully developed and deployed. Furthermore, besides time and frequency, a third dimension, space, is explored in multi-input multi-output (MIMO) systems to achieve a more reliable and higher capacity communication link. However, the positive feedback between technologies and markets always push us further. Scarce spectrum resources and nonstop market queries bring UWB technology back to the research stage again. In 2002, the Federal Communications Commissions (FCC) allocated the $3.1 - 10.6$ GHz spectrum under regulations for UWB technologies, which promoted a big boost in this area. According to the FCC, the UWB signal is any signal whose bandwidth $B \geq 500$ MHz or whose fractional bandwidth $B_f \geq 20\%$, where $B_f = 2(f_H - f_L)/(f_H + f_L)$, and f_H and f_L are the upper and lower frequency of the -10 dB emission points, respectively. This new definition categorizes UWB systems as IR-UWB systems, where each transmitted pulse instantaneously occupies the ultra-wide bandwidth, and other UWB systems, where an aggregation of at least 500 MHz bandwidth of narrowband carriers is employed, e.g., multiband OFDM systems. The UWB systems promise very high data rate and more efficient spectrum use.

UWB technology is also not new for localization. UWB radars were mainly

used for military positioning and tracking since the 1960s. In general, localization consists of two steps: first, metrics bearing location information are measured, the so-called ranging or bearing, and second, the positions are estimated based on those metrics, the so-called location information fusion. There are mainly four metrics: time-of-arrival (TOA) or time-of-flight (TOF) [11], time-difference-of-arrival (TDOA) [6, 12], angle-of-arrival (AOA) [13], and received signal strength (RSS) [14]. The ranging methods using RSS can be implemented by energy detectors, but they can only achieve a coarse resolution. Antenna arrays are required for AOA-based methods, which encumbers their popularity. On the other hand, high accuracy and potentially low cost implementation make TOA or TDOA based on IR-UWB a promising ranging method [1].

The celebrated Shannon's law ($C = B \log_2(1 + \text{SNR})$, where C is the channel capacity, B is the bandwidth of the signal, and SNR is the signal-to-noise ratio) indicates a linear relation between the channel capacity and the signal bandwidth. Meanwhile, for an additive white Gaussian noise (AWGN) channel, the ranging accuracy based on TOA estimation is known as $\sqrt{\text{Var}(\hat{d})} \geq c/(2\pi B_e \sqrt{2\text{SNR}})$ [1], where c is the signal propagation speed, and $B_e = \left(\int_{-\infty}^{\infty} f^2 |S(f)|^2 df \big/ \int_{-\infty}^{\infty} |S(f)|^2 df \right)^{1/2}$ is the effective signal bandwidth with $S(f)$ the Fourier transform of the transmitted signal. Hence, increasing the bandwidth of the signal is a very efficient way to obtain a higher channel capacity and ranging accuracy. That is the first reason why we would like to use UWB signals. Furthermore, it has a wealth of advantages, such as resolvable multipath components, harmless overlay with existing frequency allocations, low duty cycle, low probability of interception, and high probability of penetration. All these make UWB-IRs ideal for communication and localization in WSNs[*].

1.2.2 Challenges

Huge bandwidth is a double-edged sword. UWB-IRs require a Nyquist sampling rate of several tens of GHz, which is prohibitively high for practical implementation [15, 16] and conflicts with the stringent constraints of WSNs. Compressive sampling (CS) could be a rescue for UWB-IRs [17–19], which are sparse in time. CS is based on the concept that a sparse signal can be recovered from a relatively small number of measurements [20, 21]. However, the replacement of analog-to-digital convertors (ADCs) with analog-to-information convertors (AICs)[†] is not a clear answer to the question whether we really get rid of the Nyquist rate in such systems. Thus, we face the first challenge: **how to avoid the Nyquist rate sampling for UWB communications and ranging?**

---

[*]Different WSN applications require different data rates. For example, low-rate WSNs are enough for personal health monitoring, but high-rate WSNs are in need for real-time processing control loops.

[†]AICs conduct nonadaptive linear operations to the analog signal and result in discrete samples.

As we mentioned before, a TOA or TDOA metric is preferred for ranging. Since TOA or TDOA measurements are time-based, clock synchronization is essential to achieve accurate localization. Synchronization is also very important for UWB communications, but not as rigorous as it is for localization. Note that ranging is carried out between two sensor nodes, and different clock components are employed at different sensor nodes. The relative clock skew between two clocks will introduce a large clock drift over the whole ranging packet and cause severe waveform mismatch, which makes it impossible to correctly detect the TOA of the first path by averaging many frames in order to enhance the SNR. Then, we face the second challenge: **how to calibrate the clock drift for ranging?**

Furthermore, even if we correctly detect the TOA of the first path and the sensor node records a timestamp correspondingly for TOF calculation, the local timestamp has to be translated into absolute time. Since the distance is the product of the absolute TOF and the signal propagation speed, it only relates to the absolute time duration, not to any local time duration. Therefore, we face the third challenge: **how to jointly consider synchronization and localization?**

Due to the progress of the deployment of WSNs, security issues are put on the table. Localization is vulnerable to many types of attacks [22–24], such as relay attack, jamming attack and compromised node attack, etc. Since we will mainly focus on time-based localization methods, time-related attacks which tightly connect with localization and synchronization have to be considered. For example, a deceitful target sensor sends fraudulent timestamps to spoof its processing time, or submits inaccurate timestamps due to its asynchronous clock or other reasons. Hence, we face the fourth challenge: **how to take the misbehavior and unreliability of sensors into account in the development of the localization approach?**

Moreover, the low-cost and low-complexity constraints for WSNs motivate us to focus on non-coherent UWB communication systems and look for closed-form algorithms for localization. Although maximum likelihood estimators (MLEs) and convex optimization are powerful tools to approach optimum performance, high computational complexity and numerous iterations make them less attractive for WSNs. Consequently, we face the fifth challenge: **how to design a non-coherent UWB communication system and closed-form localization solutions for WSNs?**

Last but not least, in practice, sensors are not static, and they can move anywhere. We need to track their locations. Due to the nonlinear relations between ranging measurements and the coordinates of the mobile target sensors, the conventional Kalman filter (KF) cannot be used. The extended Kalman filter (EKF) [25] is most widely used to linearize the non-linear model. However, the performance of the EKF is decided by how accurate the linear approximation is. Moreover, the unscented Kalman filter (UKF) and the particle filter [26] are also effective tools to deal with nonlinear models and non-Gaussian noise for tracking. However, both of them are computationally intensive. As a result, we face the sixth challenge: **how to track a moving sensor with low complexity?**

In this thesis, we try to address all these challenges in the following chapters, which are summarized in the next section.

1.3 Thesis outline and contributions

This thesis is composed of three parts: Part I UWB communications, Part II UWB ranging, and Part III Localization and tracking. Each part is organized as follows.

- **Part I UWB communications**

  In Chapter 2 [J4][C7-8][‡], a complete detection, channel estimation, synchronization and equalization scheme for a transmitted-reference (TR) UWB system is proposed. The scheme is based on a data model which admits a moderate data rate and takes both the inter-frame interference (IFI) and the inter-symbol interference (ISI) into consideration, and replaces Nyquist sampling with sub-Nyquist sampling. Moreover, the bias caused by the inter-pulse interference (IPI) in one frame is also taken into account. Based on the analysis of the stochastic properties of the received signals, several detectors are studied and evaluated. Furthermore, a data-aided two-stage synchronization strategy is proposed, which obtains sample-level timing in the range of one symbol at the first stage and then pursues symbol-level synchronization by looking for the header at the second stage. Three channel estimators are derived to achieve joint channel and timing estimates for the first stage, namely the linear minimum mean square error (LMMSE) estimator, the least squares (LS) estimator and the matched filter (MF). We check the performance of different combinations of channel estimation and equalization schemes, and try to find the best combination, that is, the one providing a good tradeoff between complexity and performance.

- **Part II UWB ranging**

  In Chapter 3 [C6], we derive the Cramér-Rao bound (CRB) for range estimation, which does not only exploit the range information in the time delay, but also in the amplitude of the received signal. This new bound is lower than the conventional CRB that only makes use of the range information in the time delay. We investigate the new bound in an additive white Gaussian noise (AWGN) channel with attenuation by employing both narrowband (NB) signals and UWB signals. For NB signals, the new bound can be 3dB lower than the conventional CRB under certain conditions. However, there is not much difference between the new bound and the conventional CRB for UWB signals. Further, shadowing effects are added into the data model. Several CRB-like bounds for range estimation are derived to take these shadowing effects into account.

[‡]References here refer to the publication results of this thesis on pp. 14

In Chapter 4 [C5], we propose a time-of-arrival (TOA) estimation scheme using UWB impulse-radio (IR), which has a low sampling rate and is robust against clock drift. Low-rate stroboscopic sampling, which can achieve an equivalent sampling rate as high as the Nyquist sampling rate, is adopted to achieve a high resolution TOA estimate by IR-UWB. Since a long preamble is required to collect sufficient data samples, the clock drift is one of the main error sources in TOA estimation with stroboscopic sampling IR-UWB systems. The clock drift refers to the time difference between two clocks due to the relative clock skew. Taking the drift into account in our system, we first obtain a maximum-likelihood (ML) estimate of the drift. Then, we investigate several kinds of TOA estimation methods, and employ the peak selection (PS) and the jump back and search forward (JBSF) method to estimate the TOA using the averaged data samples calibrated for the drift. A theoretical performance limit is obtained by the CRB for joint estimation of the clock drift ratio and the parameters of the multipath components. Simulation results corroborate that associated drift calibration significantly reduces the TOA estimation errors, and that stroboscopic sampling can achieve the same estimation resolution as Nyquist sampling.

- **Part III Localization and tracking**

  Throughout a WSN, there are always a few sensors labeled as anchors, whose positions are known (by a system administrator during startup), and the rest are ordinary sensors, whose positions must be estimated. If the ordinary sensors can make range measurements with enough anchors, low-complexity closed-form localization estimators as proposed in the following chapters can be employed to estimate each sensor position independently. Therefore, this kind of methods is categorized as singular localization methods. However, due to the low cost and low power constraints of the sensors, their communication range is limited. Thus, not all of them can reach enough anchors. To deal with this problem, the localization can be accomplished sequentially [27, 28]. The sensors that reach enough anchors are localized first. Then, these located sensors can be viewed as new anchors that can facilitate the location estimation of other sensors. Alternatively, extra range information among sensors can be explored [2]. Sensors can cooperate with their neighbors, and their positions are jointly estimated simultaneously. This kind of methods is called aggregate localization methods. In this part, we first propose a range-based aggregate localization method in Chapter 5, and then mainly focus on time-based singular ones in the following chapters.

  In Chapter 5 [J3], we consider N sensor nodes randomly distributed in an l-dimensional space, e.g., $l = 2$ (a plane) or $l = 3$ (a space). There are no anchors. The distance measurements between nodes (given they are connected) are available, which could be obtained by TOA or RSS estimation. Note that we do not care about the approach to obtain distance measurements here. We assume that the distance measurements are composed of the true

distances corrupted by additive random noise. Our goal is to estimate the positions of all the nodes given partial pairwise distance measurements up to a translation, rotation, and reflection. Current solutions first complete the missing distances and then apply the classical multidimensional scaling (MDS) algorithm. Instead, we extend the classical MDS to a setup where the sensor network is composed of a fully connected group of nodes that communicate with each other (e.g., anchors), and a group of nodes that cannot communicate with each other, but each one of them communicates with each node in the first group. We localize the fully connected nodes by exploiting their distance measurements to the disconnected nodes. At the same time, the positions of the disconnected nodes are obtained up to a translation relative to the positions of the connected nodes. Recovering this translation, can be obtained with an additional step. Simulation results show that the proposed algorithm outperforms current MDS-like solutions to the problem.

From Chapter 6 to Chapter 8, we consider M anchor nodes and one target node, and would like to estimate or track the position of the target node. All the nodes are again distributed in an l-dimensional space. The coordinates of the anchor nodes are known and defined as $\mathbf{X}_a = [\mathbf{x}_1, \ \mathbf{x}_2, \ \ldots, \ \mathbf{x}_M]$, where the vector $\mathbf{x}_i = [x_{1,i}, \ x_{2,i}, \ \ldots, \ x_{l,i}]^T$ of length l indicates the known coordinates of the ith anchor node. We employ a vector $\mathbf{x}$ of length l to denote the unknown coordinates of the target node. Our method can easily be extended for multiple target nodes. Furthermore, since we employ TOA (or TDOA) measurements, clock synchronization is essential to achieve accurate localization. Therefore, we take clock parameters into account. We consider three different kinds of networks with different synchronization levels w.r.t. to the anchor clock. Note that these synchronization requirements are related to the anchors, and no synchronization requirements are upon the target node, which indicates that the clock of the target node can run freely. Strictly speaking, they are all asynchronous networks. Consequently, the timing relation between the target clock $C_s(t)$ and the the absolute time t can be described as [29]

$$C_s(t) \quad = \quad \alpha_s t + \theta_s, \tag{1.1}$$

where α_s and θ_s denote the unknown clock skew and clock offset of the target node clock relative to the absolute clock. To be more specific, we consider the following networks:

– Synchronous networks in Chapter 6: all the anchors are synchronized with each other. The timing relation between the ith anchor clock $C_i(t)$ and the absolute time t can be modeled as

$$C_i(t) \quad = \quad t + \theta, \ \ i = 1, 2, \ldots, M, \tag{1.2}$$

where θ is the unknown common clock offset of $C_i(t)$ relative to the absolute clock.

– Quasi-synchronous networks in Chapter 7 and Chapter 8: all the anchors have different clock offsets with each other. The clock model for the ith anchor is now modeled as

$$C_i(t) \quad = \quad t + \theta_i, \ \ i = 1, 2, \ldots, M, \qquad (1.3)$$

where θ_i is the unknown clock offset of $C_i(t)$ relative to the absolute clock.

– Fully asynchronous networks in Chapter 7: all the anchors not only have different clock offsets but also different clock skews w.r.t. each other. The clock model for the ith anchor is here given by

$$C_i(t) \quad = \quad \alpha_i t + \theta_i, \ \ i = 1, 2, \ldots, M, \qquad (1.4)$$

where α_i is the unknown clock skew of $C_i(t)$ relative to the absolute clock.

In Chapter 6 [J1], we propose several reference-free localization estimators based on TOA measurements for a scenario, where anchor nodes are synchronized, and the clock of the target node runs freely. The reference-free LS estimators that are different from the reference-based ones do not suffer from a poor reference selection. Furthermore, we generalize existing reference-based localization estimators using TOA or TDOA measurements, which are scattered over different research areas, and we shed new light on their relations. We justify that the optimal weighting matrix can compensate the influence of the reference selection for reference-based weighted LS (WLS) estimators using TOA measurements, and make all those estimators identical. However, the optimal weighting matrix cannot decouple the reference dependency for reference-based WLS estimators using a nonredundant set of TDOA measurements, but can make the estimators using the same set identical as well.

In Chapter 7 [J2][C1][C3-4], we propose robust time-based localization strategies to locate a target node with the help of anchors (nodes with known positions) in quasi-synchronous networks and fully asynchronous networks, respectively. Time-based localization approaches attract a lot of interest due to their high accuracy and potentially low cost for WSNs. However, time-based localization is tightly coupled with clock synchronization. Thus, the reliability of timestamps in time-based localization becomes an important yet challenging task to deal with. A novel ranging protocol is developed, namely asymmetric trip ranging (ATR), to reduce the communication load and explore the broadcast property of WSNs. Regardless of the reliability of the timestamp report from the target node, closed-form least-squares (LS)

estimators are derived to accurately estimate the target node position. As a result, we counter the uncertainties caused by the target node by ignoring the timestamps from this node.

In Chapter 8 [C2], we propose a Kalman filter (KF) based tracking approach to track a target node with the assistance of anchors in quasi-synchronous networks w.r.t. the synchronization level of the anchors. We again employ the asymmetric trip ranging (ATR) protocol to obtain TOA measurements and facilitate clock offset cancellation, and further derive a linear measurement model from the TOA measurements. Thus, the KF based on this linear measurement model does not have the modeling errors inherently contained in the extended Kalman filter (EKF). Furthermore, low computational complexity makes the proposed KF a promising solution for practical use.

Besides the above contributions included in this thesis, we have also developed a flexible platform shown in Fig. 1.3 in order to explore the hardware architecture of the digital receiver algorithm for the TR UWB communication system. This platform delivers a hardware experimental environment for the TR UWB scheme. It can emulate real UWB signals as inputs of our digital UWB receiver and make efficient use of a USB link between the hardware and PC. To be more specific, an arbitrary waveform generator (AWG) in the platform is programmed to generate UWB signals. These represent the transmitted TR UWB signals, convolved with a measured multipath channel impulse response and correlated by a delayed version of themselves. A dualchannel, 12-bit A/D convertor (ADC) samples the signals at 100 MHz. The data samples are transferred through a USB link from the FPGA board to the PC and demodulated by Matlab software. The function block diagram of the platform is depicted in Fig. 1.4. A digital receiver with a two-step acquisition was realized on the platform and the results were published in [C9].
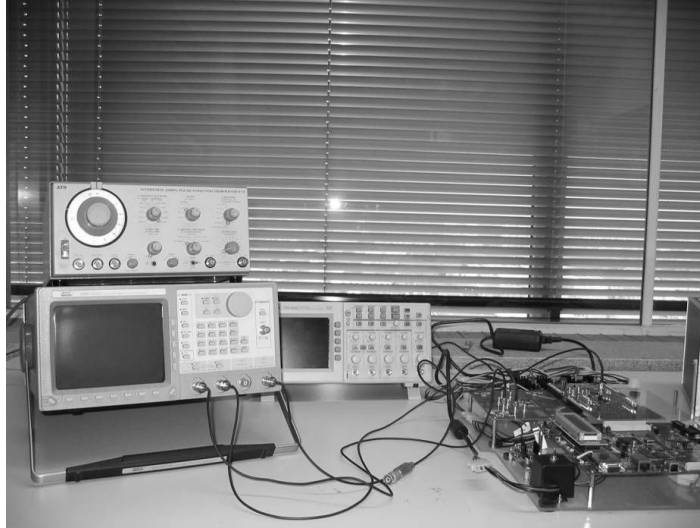
Figure 1.3: The hardware platform setup
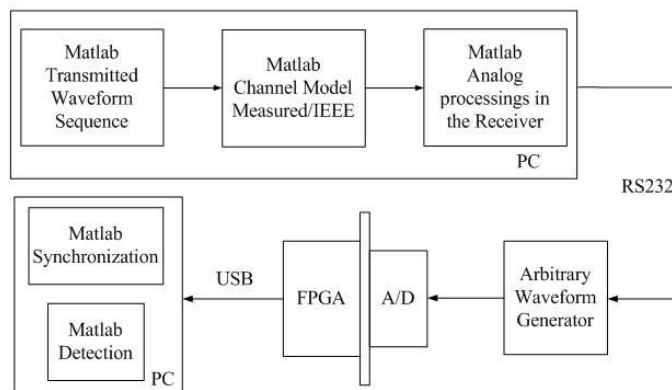
Figure 1.4: The block diagram of the hardware platform for a DHTR UWB system prototype development

- **Journal Papers:**

  [J1] Y. Wang and G. Leus, "Reference-free Time-based Localization for an Asynchronous Target", *EURASIP Journal on Advances in Signal Processing*, accepted, Sept. 2011.

  [J2] Y. Wang, X. Ma and G. Leus, "Robust Time-based Localization for

Asynchronous Networks", *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp.4397-4410, Sept. 2011.

[J3] A. Amar, Y. Wang and G. Leus, "Extending the Classical Multidimensional Scaling Algorithm Given Partial Pairwise Distance Measurements", *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 473-476, May 2010.

[J4] Y. Wang, G. Leus and A.-J. van der Veen, "Digital Receiver Design for Transmitted-Reference Ultra-Wideband Systems", *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 1-17, Feb. 2009, articleID 315264, doi:10.1155/2009/315264.

- **Conference Papers:**

[C1] Y. Wang, X. Ma and G. Leus, "Robust Time-Based Localization for Asynchronous Networks with Clock Offsets", in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, accepted, Nov. 2011.

[C2] Y. Wang, G. Leus and X. Ma, "Tracking a Mobile Node by Asynchronous Networks", in *Proc. IEEE SPAWC*, San Francisco, CA, USA, Jun. 2011.

[C3] Y. Wang, G. Leus and X. Ma, "Time-based Localization for Asynchronous Wireless Sensor Networks", in *Proc. IEEE ICASSP*, Prague, Czech Republic, pp. 3284-3287, May 2011.

[C4] Y. Wang, X. Ma and G. Leus, "A UWB Ranging-based Localization Strategy with Internal Attack Immunity", in *Proc. IEEE ICUWB*, Nanjing, China, pp. 1-4, Sept. 2010.

[C5] Y. Wang, G. Leus and H. Delic, "TOA Estimation Using UWB with Low Sampling Rate and Clock Drift Calibration", in *Proc. IEEE ICUWB*, Vancouver, Canada, pp. 612-617, Sept. 2009.

[C6] Y. Wang, G. Leus and A.J. van der Veen, "Cramer-Rao Bound for Range Estimation", in *Proc. IEEE ICASSP*, Taipei, Taiwan, pp. 3301-3304, Apr. 2009.

[C7] Y. Wang, G. Leus and A.J. van der Veen, "On Digital Receiver Design for Transmitted Reference Ultra Wideband", in *Proc. IEEE ICUWB*, Hannover, Germany, pp. 35-38, Sept. 2008.

[C8] A. Schranzhofer, Y. Wang and A.J. van der Veen, "Acquisition for a Transmitted Reference UWB Receiver", in *Proc. IEEE ICUWB*, Hannover, Germany, pp. 149-152, Sept. 2008.

[C9] Y. Wang, A. Schranzhofer, R. van Leuken and A.J. van der Veen, "Hardware Platform for Delay Hopped Transmitted Reference UWB Communication System Prototype Development", in *Proc. IEEE Pro-RISC*, Veldhoven, The Netherlands, pp. 272-275, Nov. 2007.

[C10] Y. Wang, A.J. van der Veen and R. van Leuken, "Design of a Practical Scheme for Ultra Wideband Communication", in *Proc. IEEE ISCAS*, Kos, Greek, pp. 2585-2588, May 2006.

1.4 Smart PEAS project

The research for this thesis was supported in part by the STW project DPC. 7976 "Product Quality Control Using Smart PEAS-Based UWB Technology"(Smart PEAS) within the Green and Smart Process Technologies program. The Smart PEAS project aims at studying the use of smart moving process environment actuators and sensors (PEAS) based on ultra-wideband (UWB) technology inside the process equipment for product quality control. Product quality is determined by manufacturing processes, thus processes must be rigorously characterized, understood, and controlled. Current process controls are based on measured data at a few fixed positions and carried out at a few fixed positions as well. Therefore, it is insufficient and ineffective. Smart PEAS can provide an enabling infrastructure to dramatically improve process controls in an efficient way. More measurements can be obtained to accurately characterize the process, and more controls can be preciously conducted in a distributed fashion as well.

This Smart PEAS project has been carried out by a team composed of two Ph.D. candidates, two Postdocs, one engineer and two supervisors, and was a cooperation between the Circuits & Systems (CAS) group of the Faculty of Electrical Engineering, Mathematics and Computer Sciences (EEMCS) of the Delft University of Technology and the Process & Energy (PE) group of the Faculty of Mechanical, Maritime and Materials Engineering (3ME) of the same university.

The main research work has been split into two clusters. The first cluster is related to CAS activities, whereas the second cluster contains PE tasks as follows.

Cluster 1 (CAS tasks)

1. Characterization of transmission properties of UWB-signals in various industrial media and temperatures, and related UWB antenna design.

2. Development of the UWB communications technology under conditions prevailing in chemical processes, i.e., in liquids and inside metal vessels.

3. Development of the UWB localization technology under conditions prevailing in chemical processes, i.e., in liquids and inside metal vessels.

4. Integration of the devices for communications, localization, and pro- cess variable measurements into a mobile sensor which is robust enough to be applied in a process environment.

Cluster 2 (PE tasks)

5. Defining the application scope of the new technology and identification of the measurement system. This will include the definition of e.g. sensor design

parameters/boundary conditions on geometry, chemical and temperature resistance (current maximum temperature is 150- 200C), and process safety issues.

6. Conceptual design of a measurement system in a process, including sensor removal strategies.

7. Establishing and optimizing the flow behavior of sensor geometries in process equipment. A special issue will be the interference with the normal hydrodynamic flow.

8. Process characterization by determining dynamic process parameter- position profiles which serve as validation data for advanced process models and control strategy. A first emphasis will be on positioning and temperature profiling. Depending on progress at the electronics department other measurement functionalities or actuation principles will be tested.

9. Implementation of sensors within chemical processes, with a focus on the validation of spatial distributed process models and on methods to optimize and control the local process conditions to improve the performance of the process and the product quality.

The outcome of this project consists of some fruitful results triggered by the close collaboration between the two research groups. This thesis presents some of the main research results related to the first cluster. Furthermore, together with 3UB (a TU Delft spin-off company) and Delft Engineering Services (DES), the CAS group has developed a prototype testing system within the Smart PEAS project as shown in Fig. 1.5. One of the PEAS is shown in Fig. 1.6, and consists of a ball-shaped plastic module. A bow-tie antenna is integrated on the surface, and the sensor electronics are embedded inside the module. Online localization using this platform is currently under test.

Figure 1.5: The prototype testing system for the Smart PEAS project.
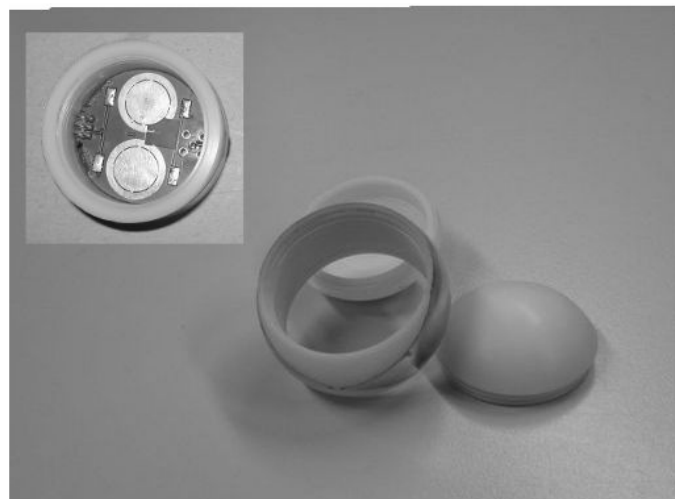
Figure 1.6: A close-up of one of the PEAS.

Part I

UWB Communications

CHAPTER 2

Transmitted-Reference Ultra-Wideband Systems

2.1   Introduction

Ultra-wideband (UWB) impulse radio (IR) techniques can provide high speed, low cost, and low complexity wireless communications with the capability to overlay existing frequency allocations [33]. Since UWB systems employ ultrashort low-duty-cycle pulses as information carriers, they suffer from stringent timing requirements [33] [34] and complex multipath channel estimation [33]. Conventional approaches require a prohibitively high sampling rate of several GHz [35] and an intensive multi-dimensional search to estimate the parameters for each multipath echo [36].

Detection, channel estimation and synchronization problems are always entangled with each other. A typical approach to address these problems is the detection-based signal acquisition [37]. A locally generated template is correlated with the received signal, and the result is compared to a threshold. How to generate a good template is the task of channel estimation, whereas how to decide the threshold is the goal of detection. Due to the multipath channel, the complexity of channel estimation grows quickly as the number of multipath components increases, and because of the fine resolution of the UWB signal, the search space is extremely large.

Recent research works on detection, channel estimation and synchronization methods for UWB have focused on low sampling rate methods [38] [39] [40] [41] or non-coherent systems, such as transmitted-reference (TR) systems [37] [42], differential detectors (DDs) [43] and energy detectors (EDs) [41] [44]. In [38], a general-

---

The results in this chapter have been published as [30–32]

ized likelihood ratio test (GLRT) for frame-level acquisition based on symbol rate sampling is proposed, which works with no or small inter-frame interference (IFI) and no inter-symbol interference (ISI). The whole training sequence is assumed to be included in the observation window without knowing the exact starting point. Due to its low duty cycle, an UWB signal belongs to the class of signals that have a finite rate of innovation [39]. Hence, it can be sampled below the Nyquist sampling rate and the timing information can be estimated by standard methods. The theory is developed under the simplest scenario and extensions are currently envisioned [45]. The timing recovery algorithm of [40] makes cross correlations of successive symbol-long received signals, in which the feedback controlled delay lines are difficult to implement. In [41], the authors address a timing estimation comparison among different types of transceivers, such as stored-reference (SR) systems, ED systems and TR systems. The ED and the TR systems belong to the class of noncoherent receivers. Although their performances are suboptimal due to the noise contaminated templates, they attract more and more interest because of their simplicity. They are also more tolerant to timing mismatches than SR systems. The algorithms in [41] are based on the assumption that the frame-level acquisition has already been achieved. Two-step strategies for acquisition are described in [46] [47]. In [46], the authors use a different search strategy in each step to speed up the procedure, which is the bit reversal search for the first step and the linear search for the second step. Meanwhile, the two-step procedure in [47] finds the block which contains the signal in the first step, and aligns with the signal at a finer resolution in the second step. Both methods are based on the assumption that coarse acquisition has already been achieved to limit the search space to the range of one frame and that there are no interferences in the signal.

From a system point of view, non-coherent receivers are considered to be more practical since they can avoid the difficulty of accurate synchronization and complicated channel estimation. One main obstacle for TR systems and DD systems is the implementation of the delay line [48]. The longer the delay line is, the more difficult it is to implement. For DD systems [43], the delay line is several frames long, whereas for TR systems, it can be only several pulses long [30], which is much shorter and easier to implement [49]. ED systems do not need a delay line, but suffer from multiple access interference [50], since they can only adopt a limited number of modulation schemes, such as on-off keying (OOK) and pulse position modulation (PPM). A two-stage acquisition scheme for TR-UWB systems is proposed in [37], which employs two sets of direct-sequence (DS) code sequences to facilitate coarse timing and fine aligning. The scheme assumes no IFI and ISI. In [51], a blind synchronization method for TR-UWB systems executes a MUSIC-kind of search in the signal subspace to achieve high resolution timing estimation. However, the complexity of the algorithm is very high because of the matrix decomposition.

Recently, a multi-user TR-UWB system that admits not only inter-pulse interference (IPI), but also IFI and ISI was proposed in [52]. The synchronization for such a system is at low-rate sample-level. The analog parts can run independently

without any feedback control from the digital parts. In this chapter, we develop a complete detection, channel estimation, synchronization and equalization scheme based on the data model modified from [52]. Moreover, the performance of different kinds of detectors is assessed. A two-stage synchronization strategy is proposed to decouple the search space and speed up synchronization. The property of the circulant matrix in the data model is exploited to reduce the computational complexity. Different combinations of channel estimators and equalizers are evaluated to find the one with the best trade-off between performance and complexity. The results confirm that the TR-UWB system is a practical scheme that can provide moderate data rate communications (e.g., in our simulation setup, the data rate is 2.2Mb/s) at a low cost.

The chapter is organized as follows. In Section 2.2, the data model presented in [52] is summarized and modified to take the unknown timing into account. Further, the statistics of the noise are derived. The detection problem is addressed in Section 2.3. Channel estimation, synchronization, and equalization are discussed in Section 2.4. Simulation results are shown and assessed in Section 2.5. Conclusions are drawn in Section 2.6.

2.2   Asynchronous single user data model

The asynchronous single user data model derived in the following paragraphs uses the data model in [52] as a starting point. We take the unknown timing into consideration and modify the model in [52].

2.2.1   Single frame

In a TR-UWB system [42] [52], pairs of pulses (doublets) are transmitted in sequence as shown in Fig.2.1. The first pulse in the doublet is the reference pulse, whereas the second one is the data pulse. Since both pulses go through the same channel, the reference pulse can be used as a "dirty template" (noise contaminated) [40] for correlation at the receiver. One frame-period T_f holds one doublet. Moreover, N_f frames constitute one symbol-period $T_s = N_f T_f$, which is carrying a symbol $s_i \in \{-1, +1\}$, spread by a pseudo random code $c_j \in \{-1, +1\}, j = 1, 2, \ldots, N_f$, which is repeatedly used for all symbols. The polarity of a data pulse is modulated by the product of a frame code and a symbol. The two pulses are separated by some delay interval D_m, which can be different for each frame. The delay intervals are in the order of nanoseconds and $D_m \ll T_f$. The receiver employs multiple correlation branches corresponding to different delay intervals. To simplify the system, we use a single delay and one correlation branch, which implies $D_m = D$. Fig. 2.1 also presents an example of the receiver structure for a single delay D. The integrate-and-dump (I&D) integrates over an interval of length T_{sam}. As a result, one frame results in $P = T_f/T_{\text{sam}}$ samples, which is assumed to be an integer.
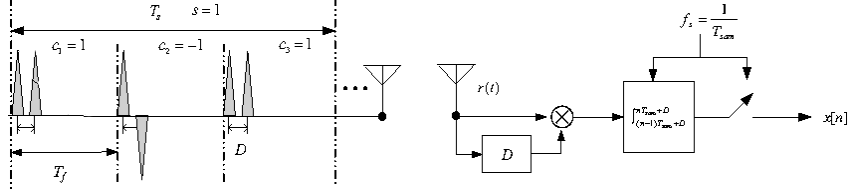
Figure 2.1: The transmitted UWB signal and the receiver structure

The received one-frame signal (jth frame of ith symbol) at the antenna output is

$$r(t) = h(t - \tau) + s_i c_j h(t - D - \tau) + n(t), \qquad (2.1)$$

where τ is the unknown timing offset, $h(t) = h_p(t) \star g(t)$ of length T_h, with $h_p(t)$ the UWB physical channel and $g(t)$ the pulse shape resulting from all the filter and antenna effects, and $n(t)$ is the bandlimited additive white Gaussian noise (AWGN) with double sided power spectral density $N_0/2$ and bandwidth B. Without loss of generality, we may assume that the unknown timing offset τ in (2.1) is in the range of one symbol period, $\tau \in [0, T_s)$, since we know the signal is present by detection at the first step (see Section 2.3) and propose to find the symbol boundary before acquiring the package header (see Section 2.4). Then, τ can be decomposed as

$$\tau = \delta \cdot T_{\text{sam}} + \epsilon, \qquad (2.2)$$

where $\delta = \lfloor \tau/T_{\text{sam}} \rfloor \in \{0, 1, \ldots, L_s - 1\}$ denotes the sample-level offset in the range of one symbol with $L_s = N_f P$, the symbol length in terms of number of samples, and $\epsilon \in [0, T_{\text{sam}})$ presents the fractional offset. Sample-level synchronization consists of estimating δ. The influence of ϵ will be absorbed in the data model and becomes invisible as we will show later.

Based on the received signal $r(t)$, the correlation branch of the receiver computes

$$
\begin{aligned}
x[n] &= \int_{(n-1)T_{\text{sam}}+D}^{nT_{\text{sam}}+D} r(t)r(t - D)dt \\
&= \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} \{[h(t - \tau) + s_i c_j h(t - D - \tau) + n(t)] \\
&\quad [h(t + D - \tau) + s_i c_j h(t - \tau) + n(t + D)]\}dt \\
&= s_i c_j \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h^2(t - \tau) + h(t - D - \tau)h(t + D - \tau)]dt \\
&\quad + \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h(t - \tau)h(t + D - \tau) + h(t - D - \tau)h(t - \tau)]dt + n_1[n],
\end{aligned}
$$

$$(2.3)$$

where

$$
\begin{aligned}
n_1[n] \;=\; n_0[n] + s_i c_j \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h(t-\tau)n(t) + h(t-D-\tau)n(t+D)]dt \\
+ \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h(t-\tau)n(t+D) + h(t+D-\tau)n(t)]dt, \qquad (2.4)
\end{aligned}
$$

with

$$
n_0[n] \;=\; \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} n(t)n(t+D)dt. \qquad (2.5)
$$

Note that $n_0[n]$ is the noise autocorrelation term, and $n_1[n]$ encompasses the signal-noise cross correlation term and the noise autocorrelation term. Their statistics will be analyzed later. Taking ϵ into consideration, we can define the channel correlation function similarly as in [52]

$$
R(\Delta, m) = \int_{(m-1)T_{\text{sam}}}^{mT_{\text{sam}}} h(t-\epsilon)h(t-\epsilon-\Delta)\, dt, \quad m = 1, 2, \ldots, \qquad (2.6)
$$

where $h(t) = 0$, when $t > T_h$ or $t < 0$. Therefore, the first term in (2.3) can be denoted as $s_i c_j \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} h^2(t-\tau)dt = s_i c_j \int_{(n-1)T_{\text{sam}}-\delta T_{\text{sam}}}^{nT_{\text{sam}}-\delta T_{\text{sam}}} h^2(t-\epsilon)dt = s_i c_j R(0, n-\delta)$. Other terms in $x[n]$ can also be rewritten in a similar way, leading $x[n]$ to be

$$
x[n] = \begin{cases} s_i c_j [R(0, n-\delta) + R(2D, n-\delta + \frac{D}{T_{\text{sam}}})] \\ +[R(D, n-\delta) + R(D, n-\delta + \frac{D}{T_{\text{sam}}})] + n_1[n], & n = \delta+1, \ldots, \delta + P_h, \\ n_0[n], & \text{elsewhere,} \end{cases}
$$
$$
\qquad (2.7)
$$

where $P_h = \lceil T_h/T_{\text{sam}} \rceil$ is the channel length in terms of number of samples and $R(0, m)$ is always non-negative. Although $R(2D, m + D/T_{\text{sam}})$ is always very small compared to $R(0, m)$, we do not ignore it to make the model more accurate. We also take the two bias terms into account, which are the cause of the IPI and are independent of the data symbols and the code. Now, we can define the $P_h \times 1$ channel energy vector $\mathbf{h}$ with entries h_m as

$$
h_m = R(0, m) + R(2D, m + \frac{D}{T_{\text{sam}}}), \quad m = 1, \ldots, P_h, \qquad (2.8)
$$

where $R(0, m) \geq 0$. Further, the $P_h \times 1$ bias vector $\mathbf{b}$ with entries b_m is defined as

$$
b_m = R(D, m) + R(D, m + \frac{D}{T_{\text{sam}}}), \quad m = 1, \ldots, P_h. \qquad (2.9)
$$

Note that these entries will change as a function of ϵ, although ϵ is not visible in

the data model. As we stated before, sample-level synchronization is limited to the estimation of δ. Using (2.8) and (2.9), $x[n]$ can be represented as

$$x[n] = \begin{cases} s_i c_j h_{n-\delta} + b_{n-\delta} + n_1[n], & n = \delta+1, \delta+2, \ldots, \delta+P_h, \\ n_0[n], & \text{elsewhere.} \end{cases} \qquad (2.10)$$

Now we can turn to the noise analysis. A number of papers have addressed the noise analysis for TR systems [53–56]. The noise properties are summarized here, and more details can be found in Appendix 2.A. We start by making the assumptions that $D \gg 1/B$, $T_{\text{sam}} \gg 1/B$, and the time-bandwidth product $2BT_{\text{sam}}$ is large enough. Under these assumptions, the noise autocorrelation term $n_0[n]$ can be assumed to be a zero mean white Gaussian random variable with variance $\sigma_0^2 = N_0^2 BT_{\text{sam}}/2$. The other noise term $n_1[n]$ includes the signal-noise cross-correlation and the noise autocorrelation, and can be interpreted as a random disturbance of the received signal. Let us define two other $P_h \times 1$ channel energy vectors $\mathbf{h}'$ and $\mathbf{h}''$ with entries h'_m and h''_m to be used in the variance of $n_1[n]$

$$h'_m = R(0, m) + R(0, m - \frac{D}{T_{\text{sam}}}), \quad m = 1, \ldots, P_h, \qquad (2.11)$$

$$h''_m = R(0, m) + R(0, m + \frac{D}{T_{\text{sam}}}), \quad m = 1, \ldots, P_h. \qquad (2.12)$$

Using those definitions and under the earlier assumptions, $n_1[n]$ can also be assumed to be a zero mean Gaussian random variable with variance $(N_0/2)\left(h'_{n-\delta} + h''_{n-\delta} + 2s_i c_j b_{n-\delta}\right) + \sigma_0^2$, $n = \delta+1, \delta+2, \ldots, \delta+P_h$. This indicates that all the noise samples are uncorrelated with each other and have a different variance depending on the data symbol, the frame code, the channel correlation coefficients, and the noise level. Note that the noise model is as complicated as the signal model.

2.2.2   Multiple frames and symbols

Now let us extend the data model to multiple frames and symbols. We assume the channel length P_h is not longer than the symbol length L_s. A single symbol with timing offset τ will then spread over at most three adjacent symbol periods. Define $\mathbf{x}_k = \left[x[(k-1)L_s + 1], x[(k-1)L_s + 2], \ldots, x[kL_s] \right]^T$, which is an L_s-long sample vector. By stacking $M + N - 1$ such received sample vectors into an $ML_s \times N$ matrix

$$\mathbf{X} = \begin{bmatrix} \mathbf{x}_k & \mathbf{x}_{k+1} & \ldots & \mathbf{x}_{k+N-1} \\ \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \ldots & \mathbf{x}_{k+N} \\ \vdots & & \ldots & \vdots \\ \mathbf{x}_{k+M-1} & \mathbf{x}_{k+M} & \ldots & \mathbf{x}_{k+M+N-2} \end{bmatrix}, \qquad (2.13)$$
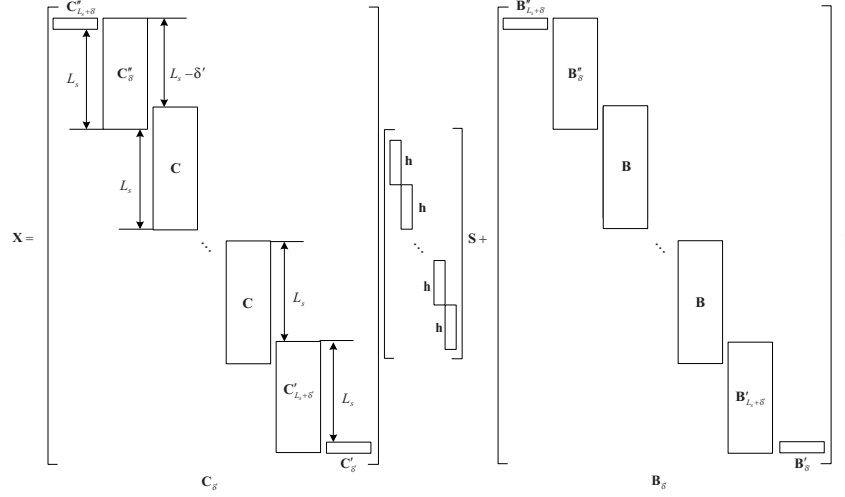
Figure 2.2: The data model structure of $\mathbf{X}$

where N indicates the number of samples in each row of $\mathbf{X}$, and M denotes the number of sample vectors in each column of $\mathbf{X}$, we obtain the following decomposition

$$\mathbf{X} = \mathbf{C}_{\delta'}(\mathbf{I}_{M+2} \otimes \mathbf{h})\mathbf{S} + \mathbf{B}_{\delta'}\mathbf{1}_{(MN_f+2N_f)\times N} + \mathbf{N}_1, \tag{2.14}$$

where $\mathbf{N}_1$ is the noise matrix similarly defined as $\mathbf{X}$,

$$\mathbf{S} \;=\; \begin{bmatrix} s_{k-1} & s_k & \cdots & s_{k+N-2} \\ s_k & s_{k+1} & \cdots & s_{k+N-1} \\ \vdots & & \cdots & \vdots \\ s_{k+M} & s_{k+M+1} & \cdots & s_{k+M+N-1} \end{bmatrix}, \tag{2.15}$$

and the structure of the other matrices is illustrated in Fig. 2.2. We first define a code matrix $\mathbf{C}$. It is a block Sylvester matrix of size $(L_s + P_h - P) \times P_h$, whose columns are shifted versions of the extended code vector: $[c_1, \mathbf{0}_{P-1}^T, c_2, \mathbf{0}_{P-1}^T, \ldots,$ $c_{N_f}, \mathbf{0}_{P-1}^T]^T$. The shift step is one sample. Its structure is shown in Fig. 2.3. The matrix $\mathbf{C}_{\delta'}$ of size $ML_s \times (MP_h + 2P_h)$ is composed of $M + 2$ block columns, where $\delta = (L_s - \delta') \bmod L_s$, $\delta' \in \{0, 1, \ldots, L_s - 1\}$. As long as there are more than two sample vectors ($M > 2$) stacked in every column of $\mathbf{X}$, the non-zero parts of the block columns will contain $M - 2$ code matrices $\mathbf{C}$. The non-zero parts of the first and last two block columns result from splitting the code matrix $\mathbf{C}$ according to δ': $\mathbf{C}'_i(2L_s - i + 1 : 2L_s, :) = \mathbf{C}(1 : i, :)$ and $\mathbf{C}''_i(1 : L_s + P_h - P - i, :) = \mathbf{C}(i+1 : L_s + P_h - P, :)$, where $\mathbf{A}(m : n, :)$ refers to column m through n of $\mathbf{A}$. The overlays between frames and symbols observed in $\mathbf{C}_{\delta'}$ indicate the existence of IFI and ISI. Then we define a bias matrix $\mathbf{B}$, which is of size $(L_s + P_h - P) \times N_f$
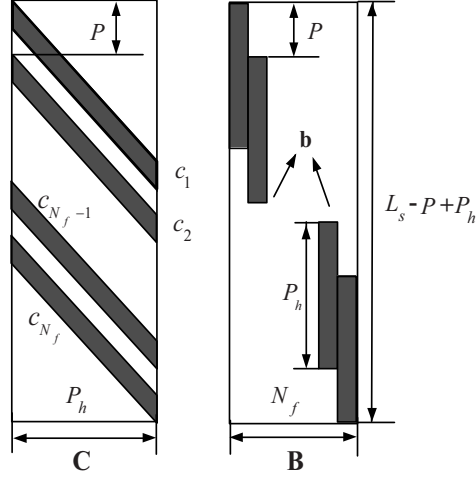
Figure 2.3: The structure of the code matrix $\mathbf{C}$ and the bias matrix $\mathbf{B}$

made up by shifted versions of the bias vector $\mathbf{b}$ with a shift step of P samples, as shown in Fig. 2.3. The matrix $\mathbf{B}_{\delta'}$ of size $ML_s \times (MN_f + 2N_f)$ also has $M+2$ block columns, the non-zero parts of which are obtained from the bias matrix $\mathbf{B}$ in the same way as $\mathbf{C}_{\delta'}$. Since the bias is independent of the data symbols and the code, it is the same for each frame. Each column of the resulting matrix $\mathbf{B}_{\delta'}\mathbf{1}_{(MN_f+2N_f)\times N}$ is the same and has a period of P samples. Defining $\mathbf{b}_f$ to be the $P \times 1$ bias vector for one such period, we have

$$\mathbf{B}_{\delta'}\mathbf{1}_{(MN_f+2N_f)\times N} = \mathbf{1}_{MN_f \times N} \otimes \mathbf{b}_f. \tag{2.16}$$

Note that $\mathbf{b}_f$ is also a function of δ, but since it is independent of the code, we cannot extract the timing information from it.

Recalling the noise analysis of the previous section, the noise matrix $\mathbf{N}_1$ has zero mean and contains uncorrelated samples with different variances. The matrix $\mathbf{\Lambda}$, which collects the variances of each element in $\mathbf{N}_1$, is

$$\begin{aligned}
\mathbf{\Lambda} &= E(\mathbf{N}_1 \odot \mathbf{N}_1) \\
&= \frac{N_0}{2}\left\{(\mathbf{H}'_{\delta'} + \mathbf{H}''_{\delta'})\mathbf{1}_{(MN_f+2N_f)\times N} + 2\mathbf{C}_{\delta'}(\mathbf{I}_{M+2} \otimes \mathbf{b})\mathbf{S}\right\} + \sigma_0^2\mathbf{1}_{ML_s \times N},
\end{aligned}$$
$$\tag{2.17}$$

where $\mathbf{H}'_{\delta'}$ and $\mathbf{H}''_{\delta'}$ have exactly the same structure as $\mathbf{B}_{\delta'}$, only using $\mathbf{h}'$ and $\mathbf{h}''$ instead of $\mathbf{b}$. They all have the same periodic property, if multiplied by $\mathbf{1}$.

Defining $\mathbf{h}'_f$ and $\mathbf{h}''_f$ to be the two $P \times 1$ vectors for one such period, we obtain

$$\mathbf{H}'_{\delta'}\mathbf{1}_{(MN_f+2N_f)\times N} \;=\; \mathbf{1}_{MN_f\times N} \otimes \mathbf{h}'_f, \qquad (2.18)$$

$$\mathbf{H}''_{\delta'}\mathbf{1}_{(MN_f+2N_f)\times N} \;=\; \mathbf{1}_{MN_f\times N} \otimes \mathbf{h}''_f. \qquad (2.19)$$

2.3 Detection

The first task of the receiver is to detect the existence of a signal. In order to separate the detection and the synchronization problems, we assume that the transmitted signal starts with a training sequence, and assign the first segment of the training sequence to detection only. In this segment, we transmit all "+1" symbols and employ all "+1" codes. It is equivalent to sending only positive pulses for some time. This kind of training sequence bypasses the code and the symbol sequence synchronization. Therefore, we do not have to consider timing issues when we handle the detection problem. The drawback is the presence of spectral peaks as a result of the periodicity. It can be solved by employing a time hopping code for the frames. We omit this in our discussion for simplicity. It is also possible to use a signal structure other than TR signals for detection, such as a positive pulse training with an ED. Although the ED doubles the noise variance due to the squaring operation, the TR system wastes half of the energy to transmit the reference pulses. Therefore, they would have a similar detection performance for the same signal-to-noise ratio (SNR), that is, the ratio of the symbol energy to the noise power spectrum density. We keep the TR structure for detection in order to avoid additional hardware for the receiver.

In the detection process, we assume that the first training segment is $2M_1$ symbols long, and the observation window is M_1 symbols long ($M_1 L_s = M_1 N_f P$ samples equivalently). We collect all the samples in the observation window, calculate a test statistic and examine whether it exceeds a threshold. If not, we jump into the next successive observation window of M_1 symbols. The $2M_1$-symbol-long training segment makes sure that there will be at least one moment, at which the M_1-symbol-long observation window is full of training symbols. In this way, we speed up our search procedure by jumping M_1 symbols. Once the threshold is exceeded, we skip the next $2M_1$ symbols in order to be out of the first segment of the training sequence and we are ready to start the channel estimation and synchronization at the sample-level (see Section 2.4). There will be situations where the observation window only partially overlaps the signal. However, for simplicity, we will not take these cases into account, when we derive the test statistic. If these cases happen and the test statistic is larger than the threshold, we declare the existence of a signal, which is true. Otherwise, we miss the detection and shift to the next observation window, which is then full of training symbols giving us a second chance to detect the signal. Therefore, we do not have to distinguish the partially overlapped cases from the overall included case. We will derive the test statistic using only these two hypotheses indicated

below. But the evaluation of the detection performance will take all the cases into account.

2.3.1   Detection problem statement

Since we only have to tell whether the whole observation window contains a signal or not, the detection problem is simplified to a binary hypothesis test. We first define the $M_1 N_f P \times 1$ sample vector $\mathbf{x} = [\mathbf{x}_k^T, \mathbf{x}_{k+1}^T, \ldots, \mathbf{x}_{k+M_1-1}^T]^T$ with entries $x[n], n = (k-1)N_f P + 1, (k-1)N_f P + 2, \ldots, (k + M_1 - 1)N_f P$, which collects all the samples in the observation window. The hypotheses are as follows

1. $\mathcal{H}_0$: there is only noise. Under $\mathcal{H}_0$, according to the analysis from the previous section, $\mathbf{x}$ is modeled as

$$\mathbf{x} = \mathbf{n}_0, \qquad (2.20)$$
$$\mathbf{x} \stackrel{a}{\sim} \mathcal{N}(0, \sigma_0^2 \mathbf{I}), \qquad (2.21)$$

   where $\mathbf{n}_0$ is the noise vector with entries $n_0[n], n = (k-1)N_f P + 1, (k-1)N_f P + 2, \ldots, (k + M_1 - 1)N_f P$ and $\stackrel{a}{\sim}$ indicates approximately distributed according to. The Gaussian approximation for $\mathbf{x}$ is valid based on the assumptions in the previous section.

2. $\mathcal{H}_1$: signal with noise is occupying the whole observation window. Under $\mathcal{H}_1$, the data model (2.14) and the noise model (2.17) can be easily specified according to the all "+1" training sequence. We define $\mathbf{H}_{\delta'}$ having the same structure as $\mathbf{B}_{\delta'}$, only taking $\mathbf{h}$ instead of $\mathbf{b}$. It also has a period of P samples in each column, if multiplied by $\mathbf{1}$. Defining $\mathbf{h}_f$ to be the $P \times 1$ vector for one such period, we have

$$\mathbf{H}_{\delta'} \mathbf{1}_{(MN_f + 2N_f) \times N} = \mathbf{1}_{MN_f \times N} \otimes \mathbf{h}_f. \qquad (2.22)$$

By selecting $M = M_1$ and $N = 1$ for (2.14), and taking (2.16), (2.18), (2.19) and (2.22) into the model, the sample vector $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = \mathbf{1}_{M_1 N_f} \otimes (\mathbf{h}_f + \mathbf{b}_f) + \mathbf{n}_1, \qquad (2.23)$$

where the zero mean noise vector $\mathbf{n}_1$ has uncorrelated entries $n_1[n], n = (k-1)N_f P + 1, (k-1)N_f P + 2, \ldots, (k + M_1 - 1)N_f P$ and the variances of each element in $\mathbf{n}_1$ are given by

$$\boldsymbol{\lambda} = E(\mathbf{n}_1 \odot \mathbf{n}_1)$$
$$= \frac{N_0}{2} \mathbf{1}_{M_1 N_f} \otimes (\mathbf{h}_f' + \mathbf{h}_f'' + 2\mathbf{b}_f) + \sigma_0^2 \mathbf{1}_{M_1 N_f P}. \qquad (2.24)$$

Due to the all "+1" training sequence, the impact of the IFI is to fold the aggregate channel response into one frame, so the frame energy remains

constant. Normally, the channel correlation function is quite narrow, so $R(D, m) \ll R(0, m)$ and $R(2D, m) \ll R(0, m)$. Then, we can have the relation

$$\mathbf{h}'_f + \mathbf{h}''_f + 2\mathbf{b}_f \approx 4(\mathbf{h}_f + \mathbf{b}_f). \qquad (2.25)$$

Defining the $P \times 1$ frame energy vector $\mathbf{z}_f = \mathbf{h}_f + \mathbf{b}_f$ with entries $z_f[i], i = 1, 2, \ldots, P$ and the frame energy $\mathcal{E}_f = \mathbf{1}_P^T \mathbf{z}_f$, we can simplify $\mathbf{x}$ and $\boldsymbol{\lambda}$

$$\mathbf{x} = \mathbf{1}_{M_1 N_f} \otimes \mathbf{z}_f + \mathbf{n}_1, \qquad (2.26)$$
$$\boldsymbol{\lambda} \approx 2N_0 \mathbf{1}_{M_1 N_f} \otimes \mathbf{z}_f + \sigma_0^2 \mathbf{1}_{M_1 N_f P}. \qquad (2.27)$$

Based on the analysis above and the assumptions from the previous section, $\mathbf{x}$ can still be assumed as a Gaussian vector in agreement with [54]

$$\mathbf{x} \overset{a}{\sim} \mathcal{N}(\mathbf{1}_{M_1 N_f} \otimes \mathbf{z}_f, \text{diag}(\boldsymbol{\lambda})), \qquad (2.28)$$

where $\text{diag}(\mathbf{a})$ indicates a square matrix with $\mathbf{a}$ on the main diagonal and zeros elsewhere.

2.3.2 Detector derivation

The test statistic is derived using $\mathcal{H}_0$ and $\mathcal{H}_1$. It is suboptimal, since it ignores other cases. But it is still useful as we have analyzed before. The *Neyman-Pearson* (NP) detector [57] decides $\mathcal{H}_1$ if

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \gamma, \qquad (2.29)$$

where γ is found by making the probability of false alarm P_{FA} to satisfy

$$P_{FA} = Pr\{L(\mathbf{x}) > \gamma; \mathcal{H}_0\} = \alpha. \qquad (2.30)$$

The test statistic is derived by taking the stochastic properties of $\mathbf{x}$ under the two hypotheses into $L(\mathbf{x})$ (2.29) and eliminating constant values. It is given by

$$T(\mathbf{x}) = \sum_{i=1}^{P} \frac{z_f[i]}{\sigma_1^2[i]} \left\{ \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} \left(x[nP+i] + \frac{N_0}{\sigma_0^2} x^2[nP+i] \right) \right\}, \qquad (2.31)$$

where $\sigma_1^2[i] = 2N_0 z_f[i] + \sigma_0^2$. A detailed derivation is presented in Appendix 2.B. Then the threshold γ will be found to satisfy

$$P_{FA} = Pr\{T(\mathbf{x}) > \gamma; \mathcal{H}_0\} = \alpha. \qquad (2.32)$$

Hence, for each observation window, we calculate the test statistic $T(\mathbf{x})$ and compare it with the threshold γ. If the threshold is exceeded, we announce that a signal is detected.

The test statistic not only depends on the noise knowledge σ_0^2, but also on the composite channel energy profile $z_f[i]$. All data samples make a weighted contribution to the test statistic, since they have different means and variances. The larger $z_f[i]/\sigma_0^2$ is, the heavier the weighting coefficient is. If we would like to employ $T(\mathbf{x})$, we have to know σ_0^2 and $z_f[i]$ first. Note that σ_0^2 can be easily estimated, when there is no signal transmitted. However, the estimation of the composite channel energy profile $z_f[i]$ is not as easy, since it appears in both the mean and the variance of $\mathbf{x}$ under $\mathcal{H}_1$.

2.3.3   Detection performance evaluation

Until now, the optimal detector for the earlier binary hypothesis test has been derived. The performance of this detector working under real circumstances has to be evaluated by taking all the cases into account. As we have described before, there are moments where the observation window partially overlays the signal. They can be modeled as other hypotheses $\mathcal{H}_j, j = 2, \ldots, M_1 N_f P$. Applying the same test statistic $T(\mathbf{x})$ under these hypotheses including $\mathcal{H}_1$, the probability of detection is defined as

$$P_{D,j} = Pr\{T(\mathbf{x}) > \gamma; \mathcal{H}_j\}, j = 1, \ldots, M_1 N_f P. \tag{2.33}$$

We would obtain $P_{D,1} > P_{D,j}, j = 2, \ldots, M_1 N_f P$. Since the observation window collects the maximum signal energy under $\mathcal{H}_1$ and the test statistic is optimal to detect $\mathcal{H}_1$, it should have the highest possibility to detect the signal. Furthermore, if we miss the detection under $\mathcal{H}_j, j = 1, \ldots, M_1 N_f P$, we still have a second chance to detect the signal with a probability of $P_{D,1}$ in the next observation window, recalling that the training sequence is $2M_1$ symbols long. Therefore, the total probability of detection for this testing procedure is $P_{D,j} + (1 - P_{D,j})P_{D,1}, j = 1, \ldots, M_1 N_f P$, which is larger than $P_{D,1}$ and not larger than $P_{D,1} + (1 - P_{D,1})P_{D,1}$. Since all hypotheses $\mathcal{H}_j, j = 1, \ldots, M_1 N_f P$ have equal probability, we can obtain that the overall probability of detection P_{D_o} for the detector $T(\mathbf{x})$ as

$$P_{D_o} = \frac{1}{M_1 N_f P} \sum_{j=1}^{M_1 N_f P} \{P_{D,j} + (1 - P_{D,j})P_{D,1}\}, j = 1, \ldots, M_1 N_f P. \tag{2.34}$$

where $P_{D,1} < P_{D_o} < P_{D,1} + (1 - P_{D,1})P_{D,1}$. Since the analytical evaluation of P_{D_o} is very complicated, we just derive the theoretical performance of $P_{D,1}$ under $\mathcal{H}_1$. In the simulations section, we will obtain the total P_{D_o} by Monte Carlo simulations and compare it with $P_{D,1}$ and $P_{D,1} + (1 - P_{D,1})P_{D,1}$, which can be

used as boundaries for P_{D_o}.

A theoretical evaluation of $P_{D,1}$ is carried out by first analyzing the stochastic properties of $T(\mathbf{x})$. As $T(\mathbf{x})$ is composed of two parts, we can define

$$T_1(\mathbf{x}) = \sum_{i=1}^{P} \frac{z_f[i]}{\sigma_1^2[i]} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x[nP+i], \qquad (2.35)$$

$$T_2(\mathbf{x}) = \sum_{i=1}^{P} \frac{z_f[i]}{\sigma_1^2[i]} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i]. \qquad (2.36)$$

Then we have

$$T(\mathbf{x}) = T_1(\mathbf{x}) + \frac{N_0}{\sigma_0^2} T_2(\mathbf{x}). \qquad (2.37)$$

First, we have to know the probability density function (PDF) of $T(\mathbf{x})$. However, due to the correlation between the two parts, it can only be found in an empirical way by generating enough samples of $T(\mathbf{x})$ and making a histogram to depict the relative frequencies of the sample ranges. Therefore, we simply assume that $T_1(\mathbf{x})$ and $T_2(\mathbf{x})$ are uncorrelated, and $T(\mathbf{x})$ is a Gaussian random variable. The mean (variance) of $T(\mathbf{x})$ is the sum of the weighted means (variances) of the two parts. The larger the sample number $M_1 N_f P$ is, the better the approximation is, but also the longer the detection time is. There is a tradeoff. In summary, $T(\mathbf{x})$ follows a Gaussian distribution as follows

$$T(\mathbf{x}) \overset{a}{\sim} \mathcal{N}\left(E(T_1(\mathbf{x})) + \frac{N_0}{\sigma_0^2} E(T_2(\mathbf{x})), \text{var}(T_1(\mathbf{x})) + \frac{N_0^2}{\sigma_0^4} \text{var}(T_2(\mathbf{x}))\right). \qquad (2.38)$$

The mean and the variance of $T_1(\mathbf{x})$ can be easily obtained based on the assumption that $\mathbf{x}$ is a Gaussian vector. The stochastic properties of $T_2(\mathbf{x})$ are much more complicated. More details are discussed in Appendix 2.C. All the performance approximations are summarized in Table 2.1, where the function $Q(\cdot)$ is the right-tail probability function for a Gaussian distribution.

A special case occurs when $P = 1$, which means that one sample is taken per frame ($T_{\text{sam}} = T_f$). For this case where no oversampling is used, we have constant energy $\mathcal{E}_f$ and constant noise variance $\sigma_1^2 = 2N_0 \mathcal{E}_f + \sigma_0^2$ for each frame. Then the weighting parameters for each sample in the detector would be exactly the same.

We can eliminate them and simplify the test statistic to

$$T_1'(\mathbf{x}) = \sum_{n=(k-1)N_f+1}^{(k+M_1-1)N_f} x[n], \qquad (2.39)$$

$$T_2'(\mathbf{x}) = \sum_{n=(k-1)N_f+1}^{(k+M_1-1)N_f} x^2[n], \qquad (2.40)$$

and

$$T'(\mathbf{x}) = T_1'(\mathbf{x}) + \frac{N_0}{\sigma_0^2} T_2'(\mathbf{x}). \qquad (2.41)$$

Therefore, $T_2'(\mathbf{x})/\sigma_0^2$ will follow a central Chi-squared distribution under $\mathcal{H}_0$ and $T_2'(\mathbf{x})/\sigma_1^2$ will follow a noncentral Chi-squared distribution under $\mathcal{H}_1$. We calculate the threshold for $T_2'(\mathbf{x})$ as

$$\gamma_2' = \sigma_0{}^2 Q_{\chi^2_{M_1 N_f}}^{-1}(\alpha), \qquad (2.42)$$

and the probability of detection under $\mathcal{H}_1$ as

$$P_{D,1} = Q_{\chi^2_{M_1 N_f}\left(M_1 N_f \mathcal{E}_f^2/\sigma_1^2\right)}\left(\frac{\gamma_2'}{\sigma_1^2}\right), \qquad (2.43)$$

Table 2.1: Statistical Analysis and Performance Evaluation for Different Detectors, $P > 1, T_{\mathrm{sam}} = T_f/P$

		$T_1(\mathbf{x})$	$T_2(\mathbf{x})$	$T(\mathbf{x})$
$\mathcal{H}_0$	μ	$\mu_{T_{1,0}} = 0$	$\mu_{T_{2,0}} = M_1 N_f \sigma_0{}^2 \sum_{i=1}^{P} \dfrac{z_f[i]}{\sigma_1^2[i]}$	$\mu_{T_0} = \mu_{T_{1,0}} + \dfrac{N_0}{\sigma_0^2}\mu_{T_{2,0}}$
	σ^2	$\sigma^2_{T_{1,0}} = M_1 N_f \sigma_0{}^2 \sum_{i=1}^{P} \dfrac{z_f^2[i]}{\sigma_1^4[i]}$	$\sigma^2_{T_{2,0}} = 2M_1 N_f \sigma_0{}^4 \sum_{i=1}^{P} \dfrac{z_f^2[i]}{\sigma_1^4[i]}$	$\sigma^2_{T_0} = \sigma^2_{T_{1,0}} + \dfrac{N_0^2}{\sigma_0^4}\sigma^2_{T_{2,0}}$
$\mathcal{H}_1$	μ	$\mu_{T_{1,1}} = M_1 N_f \sum_{i=1}^{P} \dfrac{z_f^2[i]}{\sigma_1^2[i]}$	$\mu_{T_{2,1}} = M_1 N_f \sum_{i=1}^{P} z_f[i]\left(1 + \dfrac{z_f^2[i]}{\sigma_1^2[i]}\right)$	$\mu_{T_1} = \mu_{T_{1,1}} + \dfrac{N_0}{\sigma_0^2}\mu_{T_{2,1}}$
	σ^2	$\sigma^2_{T_{1,1}} = M_1 N_f \sum_{i=1}^{P} \dfrac{z_f^2[i]}{\sigma_1^2[i]}$	$\sigma^2_{T_{2,1}} = 2M_1 N_f \sum_{i=1}^{P} z_f^2[i]\left(1 + \dfrac{2z_f^2[i]}{\sigma_1^2[i]}\right)$	$\sigma^2_{T_1} = \sigma^2_{T_{1,1}} + \dfrac{N_0^2}{\sigma_0^4}\sigma^2_{T_{2,1}}$
P_{FA}		$Q\left(\dfrac{\gamma_1}{\sigma_{T_{1,0}}}\right) = \alpha$	$Q\left(\dfrac{\gamma - \mu_{T_{2,0}}}{\sigma_{T_{2,0}}}\right) = \alpha$	$Q\left(\dfrac{\gamma - \mu_{T_0}}{\sigma_{T_0}}\right) = \alpha$
γ		$\gamma_1 = \sigma_{T_{1,0}} Q^{-1}(\alpha)$	$\gamma_2 = \sigma_{T_{2,0}} Q^{-1}(\alpha) + \mu_{T_{2,0}}$	$\gamma = \sigma_{T_0} Q^{-1}(\alpha) + \mu_{T_0}$
$P_{D,1}$		$Q\left(\dfrac{\gamma_1 - \mu_{T_{1,1}}}{\sigma_{T_{1,1}}}\right)$	$Q\left(\dfrac{\gamma_2 - \mu_{T_{2,1}}}{\sigma_{T_{2,1}}}\right)$	$Q\left(\dfrac{\gamma - \mu_{T_1}}{\sigma_{T_1}}\right)$
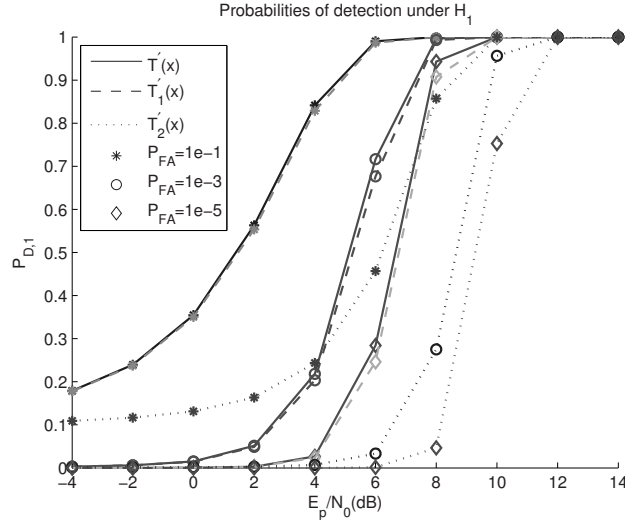
Figure 2.4: Performance comparison between $T'(\mathbf{x})$ and its components $T'_1(\mathbf{x})$ and $T'_2(\mathbf{x})$

where the functions $Q_{\chi^2_\nu}(x)$ and $Q_{\chi^2_\nu(\lambda)}(x)$ are the right-tail probability functions for a central and noncentral Chi-squared distribution, respectively. The statistics of $T'_1(\mathbf{x})$ can be obtained by taking $P = 1$, $z_f[i] = \mathcal{E}_f$ and $\sigma^2_1[i] = \sigma^2_1$ into Table 2.1, and multiplying the means with $\sigma^2_1/\mathcal{E}_f$ and the variances with $\sigma^4_1/\mathcal{E}^2_f$. As a result, the threshold γ'_1 for $T'_1(\mathbf{x})$ is $\sqrt{M_1 N_f \sigma^2_0} Q^{-1}(\alpha)$, which can be easily obtained. The $P_{D,1}$ of $T'(\mathbf{x})$ could be evaluated in the same way as $T(\mathbf{x})$ by Table 2.1.

The theoretical contributions of $T'_1(\mathbf{x})$ and $T'_2(\mathbf{x})$ to $T'(\mathbf{x})$ are assessed in Fig. 2.4. The simulation parameters are set to $M_1 = 8$, $N_f = 15$, $T_f = 30$ns, $T_p = 0.2$ns and $B \approx 2/T_p$. For the definition of E_p/N_0 we refer to Section 2.5. The detector based on $T'_1(\mathbf{x})$ (dashed lines) plays a key role in the performance of the detector based on $T'(\mathbf{x})$ (solid lines) under $\mathcal{H}_1$. For low SNR, they are almost the same, since $T'_1(\mathbf{x})$ can be directly derived by ignoring the signal-noise cross correlation term in the noise variance under $\mathcal{H}_1$. There is a small difference between them for medium SNRs. $T'_2(\mathbf{x})$ (dotted lines) has a performance loss of about 4dB compared to $T'(\mathbf{x})$. Thanks to the ultra-wide bandwidth of the signal, the weighting parameter $N_0/\sigma_0{}^2$ greatly reduces the influence of $T'_2(\mathbf{x})$ on $T'(\mathbf{x})$. It enhances the performance of $T'(\mathbf{x})$ slightly in the medium SNR range. According to these simulation results and the impact of the weighting parameter N_0/σ^2_0, we can employ $T'_1(\mathbf{x})$ instead of $T'(\mathbf{x})$. It has a much lower calculation cost and almost the same performance as $T'(\mathbf{x})$.

Furthermore, the influence of the oversampling rate P to the $P_{D,1}$ of $T(\mathbf{x})$

can be ignored, because the oversampling only affects the performance of $T_2(\mathbf{x})$, which only has a very small influence on $T(\mathbf{x})$. Therefore, the impact of the oversampling can be neglected. In Section 2.5, we will evaluate the $P_{D,1}$ of $T(\mathbf{x})$ using the IEEE UWB channel model by a quasi-analytical method and also by Monte Carlo simulations. Based on the simulation results in this section, we can predict that for small P ($P > 1$) the $P_{D,1}$ for $T(\mathbf{x})$ will be more or less the same as the $P_{D,1}$ for $T'(\mathbf{x})$ or $T_1'(\mathbf{x})$.

2.4 Channel estimation, synchronization and equalization

After successful signal detection, we can start the channel estimation and synchronization phase. The sample-level synchronization finds out the symbol boundary (estimates the unknown offset δ), and the result can later on be used for symbol-level synchronization to acquire the header. This two-stage synchronization strategy decomposes a two-dimensional search into two one-dimensional searches, reducing the complexity. The channel estimates and the timing information can be used for the equalizer construction. Finally, the demodulated symbols can be obtained.

2.4.1 Channel estimation

Bias estimation

As we have seen in the asynchronous data model, the bias term is undesired. It does not have any useful information, but disturbs the signal. We will show that this bias seriously degrades the channel estimation performance later on. The second segment of the training sequence consists of "+1, −1" symbol pairs employing a random code. The total length of the second segment should be $M_1 + 2N_s$ symbols, which includes the budget for jumping $2M_1$ symbols after the detection. The "+1, −1" symbol pairs can be used for bias estimation as well as channel estimation. Since the bias is independent of the data symbols and the useful signal part has zero mean, due to the "+1, −1" training symbols, we can estimate the $L_s \times 1$ bias vector of one symbol, $\mathbf{b}_s = \mathbf{1}_{N_f} \otimes \mathbf{b}_f$, as

$$\hat{\mathbf{b}}_s = \frac{1}{2N_s}[\mathbf{x}_k \quad \mathbf{x}_{k+1} \quad \ldots \quad \mathbf{x}_{k+2N_s-1}]\mathbf{1}_{2N_s}. \tag{2.44}$$

Channel estimation

To take advantage of the second segment of the training sequence, we stack the data samples as

$$\widetilde{\mathbf{X}} = \left[\begin{array}{cccc} \mathbf{x}_k & \mathbf{x}_{k+2} & \ldots & \mathbf{x}_{k+2N_s-2} \\ \mathbf{x}_{k+1} & \mathbf{x}_{k+3} & \ldots & \mathbf{x}_{k+2N_s-1} \end{array} \right], \tag{2.45}$$

which is equivalent to picking only odd columns of $\mathbf{X}$ in (2.14) with $M = 2$ and $N = 2N_s - 1$. As a result, each column depends on the same symbols, which leads to a great simplification of the decomposition in (2.14)

$$
\begin{aligned}
\widetilde{\mathbf{X}} &= [(\mathbf{C}'_{L_s+\delta'} + \mathbf{C}''_{L_s+\delta'}) \quad (\mathbf{C}'_{\delta'} + \mathbf{C}''_{\delta'})](\mathbf{I}_2 \otimes \mathbf{h})[-s_k \quad s_k]^T \mathbf{1}_{N_s}^T \\
&\quad + \mathbf{1}_{2\times N_s} \otimes \mathbf{b}_s + \widetilde{\mathbf{N}}_1,
\end{aligned} \tag{2.46}
$$

where $\widetilde{\mathbf{N}}_1$ is the noise matrix similarly defined as $\widetilde{\mathbf{X}}$. For simplicity, we only count the noise autocorrelation term with zero mean and variance σ_0^2 into $\widetilde{\mathbf{N}}_1$, while σ_0^2 can be easily estimated in the absence of a signal. Because we jump into this second segment of the training sequence after detecting the signal, we do not know whether the symbol s_k is "+1" or "-1". Rewriting (2.46) in another form leads to

$$
\widetilde{\mathbf{X}} = \mathbf{C}_s \mathbf{h}_{ss\delta} \mathbf{1}_{N_s}^T + \mathbf{1}_{2\times N_s} \otimes \mathbf{b}_s + \widetilde{\mathbf{N}}_1, \tag{2.47}
$$

where $\mathbf{C}_s$ is a known $2L_s \times 2L_s$ circulant code matrix, whose first column is $[c_1, \mathbf{0}_{P-1}^T, c_2, \mathbf{0}_{P-1}^T, \ldots, c_{N_f}, \mathbf{0}_{L_s+P-1}^T]^T$, and the vector $\mathbf{h}_{ss\delta}$ of length $2L_s$ blends the timing and the channel information, which contains two channel energy vectors with different signs, $s_k \mathbf{h}$ and $-s_k \mathbf{h}$, located according to δ

$$
\mathbf{h}_{ss\delta} = \begin{cases} \text{circshift}([s_k\mathbf{h}^T, \mathbf{0}_{L_s-P_h}^T, -s_k\mathbf{h}^T, \mathbf{0}_{L_s-P_h}^T]^T, \delta), & \delta \neq 0 \\ [-s_k\mathbf{h}^T, \mathbf{0}_{L_s-P_h}^T, s_k\mathbf{h}^T, \mathbf{0}_{L_s-P_h}^T]^T, & \delta = 0 \end{cases}, \tag{2.48}
$$

where $\text{circshift}(\mathbf{a}, n)$ circularly shifts the values in the vector $\mathbf{a}$ by $|n|$ elements (down if $n > 0$ and up if $n < 0$). According to (2.47) and assuming the channel energy has been normalized, the linear minimum mean square error (LMMSE) estimate of $\mathbf{h}_{ss\delta}$ then is

$$
\hat{\mathbf{h}}_{ss\delta} = \mathbf{C}_s^H (\mathbf{C}_s \mathbf{C}_s^H + \frac{\sigma_0^2}{N_s}\mathbf{I})^{-1} \frac{1}{N_s}(\widetilde{\mathbf{X}} - \mathbf{1}_{2\times N_s} \otimes \mathbf{b}_s)\mathbf{1}_{N_s}. \tag{2.49}
$$

Defining

$$
\hat{\mathbf{h}}_{s\delta} = \frac{\left[\hat{\mathbf{h}}_{ss\delta}(1:L_s) - \hat{\mathbf{h}}_{ss\delta}(L_s+1:2L_s)\right]}{2}, \tag{2.50}
$$

where $\mathbf{a}(m:n)$ refers to element m through n of $\mathbf{a}$, we can obtain a symbol-long LMMSE channel estimate as

$$
\hat{\mathbf{h}}_\delta = |\hat{\mathbf{h}}_{s\delta}|. \tag{2.51}
$$

According to a property of circulant matrices, $\mathbf{C}_s$ can be decomposed as $\mathbf{C}_s = \mathcal{F}\boldsymbol{\Omega}\mathcal{F}^H$, where $\mathcal{F}$ is the normalized DFT matrix of size $2L_s \times 2L_s$ and $\boldsymbol{\Omega}$ is a diagonal matrix with the frequency components of the first row of $\mathbf{C}_s$ on the diagonal. Hence, the matrix inversion in (2.49) can be simplified dramatically. Observing that $\mathbf{C}_s^H(\mathbf{C}_s\mathbf{C}_s^H + (\sigma_0^2/N_s)\mathbf{I})^{-1}$ is a circulant matrix, the bias term
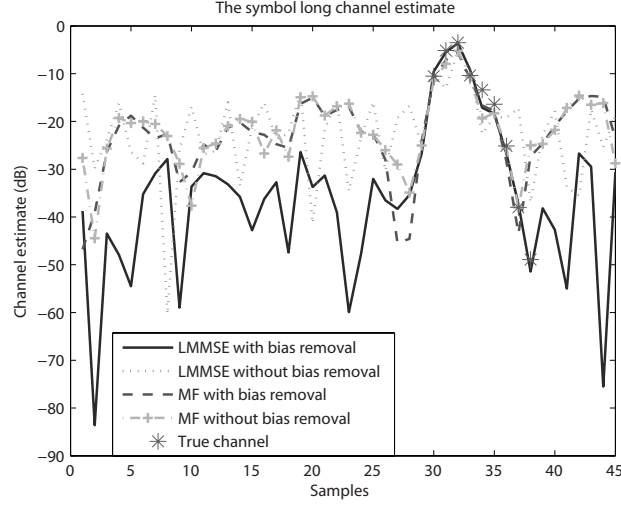
Figure 2.5: The symbol-long channel estimate $\hat{\mathbf{h}}_\delta$ with bias removal and $|\hat{\mathbf{h}}_{ss\delta}(1:L_s)|$ without bias removal, when SNR is 18 dB

actually does not have to be removed in (2.49), since it is implicitly removed when we calculate (2.50). Therefore, we do not have to estimate the bias term explicitly for channel estimation and synchronization.

When the SNR is high, $\parallel \mathbf{C}_s\mathbf{C}_s^H \parallel_F \gg \parallel (\sigma_0^2/N_s)\mathbf{I} \parallel_F$, (2.49) can be replaced by

$$\hat{\mathbf{h}}_{ss\delta} = \frac{1}{N_s}\mathcal{F}\boldsymbol{\Omega}^{-1}\mathcal{F}^H\big(\widetilde{\mathbf{X}} - \mathbf{1}_{2\times N_s} \otimes \mathbf{b}_s\big)\mathbf{1}_{N_s}. \tag{2.52}$$

It is a least squares (LS) estimator and equivalent to a deconvolution of the code sequence in the frequency domain. On the other hand, when the SNR is low, $\parallel \mathbf{C}_s\mathbf{C}_s^H \parallel_F \ll \parallel (\sigma_0^2/N_s)\mathbf{I} \parallel_F$, (2.49) becomes

$$\hat{\mathbf{h}}_{ss\delta} = \frac{1}{\sigma_0^2}\mathcal{F}\boldsymbol{\Omega}^H\mathcal{F}^H\big(\widetilde{\mathbf{X}} - \mathbf{1}_{2\times N_s} \otimes \mathbf{b}_s\big)\mathbf{1}_{N_s}, \tag{2.53}$$

which is equivalent to a matched filter (MF). The MF can also be processed in the frequency domain. The LMMSE estimator in (2.49), the LS estimator in (2.52) and the MF in (2.53) all have a similar computational complexity. However for the LMMSE estimator, we have to estimate σ_0^2 and the channel energy.

As an example, we show the performance of these channel estimates under high SNR conditions (the simulation parameters can be found in Section 2.5). Fig. 2.5 indicates the symbol-long channel estimate $\hat{\mathbf{h}}_\delta$ with bias removal and $|\hat{\mathbf{h}}_{ss\delta}(1:L_s)|$

without bias removal, where $\hat{\mathbf{h}}_{ss\delta} = \mathbf{C}_s^H(\mathbf{C}_s\mathbf{C}_s^H + (\sigma_0^2/N_s)\mathbf{I})^{-1}(1/N_s)\widetilde{\mathbf{X}}\mathbf{1}_{N_s}$ for the LMMSE and $\hat{\mathbf{h}}_{ss\delta} = (1/\sigma_0^2)\mathcal{F}\mathbf{\Omega}^H\mathcal{F}^H\widetilde{\mathbf{X}}\mathbf{1}_{N_s}$ for the MF. When the SNR is high, the LMMSE estimator is expected to have a similar performance as the LS estimator. Thus, we omit the LS estimator in Fig. 2.5. The MF for $\hat{\mathbf{h}}_\delta$ (dashed line) has a higher noise floor than the LMMSE estimator for $\hat{\mathbf{h}}_\delta$ (solid line), since its output is the correlation of the channel energy vector with the code autocorrelation function. The bias term lifts the noise floor of the channel estimate resulting from the LMMSE estimator (dotted line) and distorts the estimation, while it does not have much influence on the MF (dashed line with + markers). The stars in the figure present the real channel parameters as a reference. The position of the highest peak for each curve in Fig. 2.5 indicates the timing information and the area around this highest peak is the most interesting part, since it shows the estimated channel energy profile. Although the LMMSE estimator without bias suppresses the estimation errors over the whole symbol period, it has a similar performance as all the other estimators in the interesting part.

2.4.2   Sample-level synchronization

The channel estimate $\hat{\mathbf{h}}_\delta$ has a duration of one symbol. But we know that the true channel will generally be much shorter than the symbol period. We would like to detect the part that contains most of the channel energy, and cut out the other part in order to be robust against noise. This basically means that we have to estimate the unknown timing δ. Define the search window length as L_w in terms of the number of samples ($L_w > 1$). The optimal length of the search window depends on the channel energy profile and the SNR. We will show the impact of different window lengths on the estimation of δ in the next section. Define $\hat{\mathbf{h}}_{w\delta} = [\hat{\mathbf{h}}_{s\delta}^T, -\hat{\mathbf{h}}_{s\delta}^T(1 : L_w - 1)]^T$, and define $\hat{\delta}$ as the δ estimate

$$\hat{\delta} = \underset{\delta}{\operatorname{argmax}} |\sum_{n=\delta+1}^{\delta+L_w} \hat{\mathbf{h}}_{w\delta}(n)|. \tag{2.54}$$

This is motivated as follows. According to the definition of $\hat{\mathbf{h}}_{s\delta}$, when $\delta > L_s - P_h$, $\hat{\mathbf{h}}_{s\delta}$ will contain channel information partially from $s_k\mathbf{h}$ and partially from $-s_k\mathbf{h}$, which have opposite signs. In order to estimate δ, we circularly shift the search window to check all the possible sample positions in $\hat{\mathbf{h}}_{s\delta}$ and find the position where the search window contains the maximum energy. If we do not adjust the signs of the two parts, the δ estimation will be incorrect when the real δ is larger than $L_s - P_h$. This is because the two parts will cancel each other, when both of them are encompassed by the search window. That is the reason why we construct $\hat{\mathbf{h}}_{w\delta}$ by inverting the sign of the first $L_w - 1$ samples in $\hat{\mathbf{h}}_{s\delta}$ and attaching them to the end of $\hat{\mathbf{h}}_{s\delta}$. Moreover, the estimator (2.54) benefits from averaging the noise before taking the absolute value.

2.4.3   Equalization and symbol-level synchronization

Based on the channel estimate $\hat{\mathbf{h}}_\delta$ and the timing estimate $\hat{\delta}$, we select a part of $\hat{\mathbf{h}}_\delta$ to build three different kinds of equalizers. Since the MF equalizer cannot handle IFI and ISI, we only select the first P samples (the frame length in terms of number of samples) of $\mathrm{circshift}(\hat{\mathbf{h}}_\delta, -\hat{\delta})$ as $\hat{\mathbf{h}}_p$. The code matrix $\mathbf{C}$ is specified by assigning $P_h = P$. The estimated bias $\hat{\mathbf{b}}_s$ can be used here. We skip the first $\hat{\delta}$ data samples and collect the rest data samples in a matrix $\mathbf{X}_{\hat{\delta}}$ of size $L_s \times N$ as in the data model (2.14) but with $M = 1$. Therefore, the MF equalizer is constructed as

$$\hat{\mathbf{s}}^T = \mathrm{sign}\{(\mathbf{C}\hat{\mathbf{h}}_p)^T(\mathbf{X}_{\hat{\delta}} - \mathbf{1}_{1\times N} \otimes \hat{\mathbf{b}}_s)\}, \tag{2.55}$$

where $\hat{\mathbf{s}}$ is the estimated symbol vector. Moreover, we also construct a zero-forcing (ZF) equalizer and an LMMSE equalizer by replacing $\mathbf{h}$ with $\hat{\mathbf{h}}$, which collects the first $\hat{P}_h$ samples (the channel length estimate in terms of number of samples) of $\mathrm{circshift}(\hat{\mathbf{h}}_\delta, -\hat{\delta})$, and using $\hat{\delta}' = (L_s - \hat{\delta}) \bmod L_s$ in the data model (2.14). The channel length estimate $\hat{P}_h$ could be obtained by setting a threshold (e.g., 10% of the maximum value of $\hat{\mathbf{h}}_\delta$) and counting the number of samples beyond it in $\hat{\mathbf{h}}_\delta$. These equalizers can resolve the IFI and the ISI to achieve a better performance at the expense of a higher computational complexity. The estimated bias $\hat{\mathbf{b}}_s$ can also be used. We collect the samples in a data matrix $\mathbf{X}$ of size $2L_s \times N$ similar as the data model (2.14) with $M = 2$. Then the ZF equalizer gives

$$\hat{\mathbf{S}} = \mathrm{sign}\{(\mathbf{C}_{\hat{\delta}'}(\mathbf{I}_4 \otimes \hat{\mathbf{h}}))^\dagger(\mathbf{X} - \mathbf{1}_{2\times N} \otimes \hat{\mathbf{b}}_s)\}, \tag{2.56}$$

and the LMMSE equalizer gives

$$\hat{\mathbf{S}} = \mathrm{sign}\{(\hat{\mathbf{\Phi}}^H\hat{\mathbf{\Phi}} + \sigma_0^2\mathbf{I}_4)^{-1}\hat{\mathbf{\Phi}}^H(\mathbf{X} - \mathbf{1}_{2\times N} \otimes \hat{\mathbf{b}}_s)\}, \tag{2.57}$$

where $\hat{\mathbf{\Phi}} = \mathbf{C}_{\hat{\delta}'}(\mathbf{I}_4 \otimes \hat{\mathbf{h}})$. $\hat{\mathbf{S}}$ is a $4 \times N$ symbol matrix. We can choose either the second or the third row of $\hat{\mathbf{S}}$ as the demodulated symbol sequence.

Until now, the sample-level synchronization confirms the boundaries of the symbols. However it is not able to explore the boundary of the training header, since the second segment of the training sequence just employs pairs of "+1,-1" symbols. After the sample-level synchronization, the demodulation is triggered. The third segment of the training sequence is a known training symbol pattern. Once we find the matching symbol pattern, we can distinguish the training header. Symbol-level synchronization is then accomplished. To summarize the training segments used in every stage, the overall structure of the training sequence is shown in Fig. 2.6.

| Segment 1, no code | Segment 2, PN code | Segment 3, the header, PN code | Data |
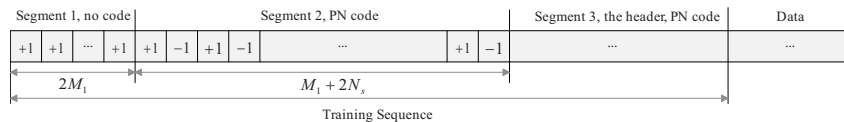
Figure 2.6: The signal structure for training sequence

2.5 Simulation results

We evaluate the performance of different detectors and the performance of different combinations of channel estimation and equalization schemes for a single user and single delay TR-UWB system. We use a Gaussian second derivative pulse that is 0.2 ns wide. The delay interval D between two pulses in a doublet is 4 ns. The first segment of the training sequence is $2M_1 = 16$ symbols long, all of which are composed of positive pulses. Hence, the observation window includes $M_1 = 8$ symbols. The second segment of the training sequence has $M_1 + 2N_s = 38$ symbols, and employs a pseudonoise (PN) code sequence. The code length N_f is 15. The frame-period T_f is 30 ns. The IEEE UWB channel model CM3 [58] is employed and truncated to 90 ns, which represents a NLOS channel. The oversampling rate P is 3, which results in $T_{\text{sam}} = 10$ ns. We define E_p/N_0 as the received aggregate pulse energy to noise ratio with $E_p = \int |h(t)|^2 dt$, where $h(t)$ represents the composite channel impulse response including pulse shaping and antenna effects as we have explained before (see Section II.A). The system sampling rate is 50GHz for Matlab simulations.

The test statistics $T(\mathbf{x})$ in (2.37) and $T_1'(\mathbf{x})$ in (2.39) are assessed in both a theoretical way by using the results in Table 2.1 and an experimental way by running Monte Carlo simulations. Fig. 2.7 shows the probability of detection $P_{D,1}$ for the test statistics. The theoretical $P_{D,1}$ of $T(\mathbf{x})$ with $P = 3$ is evaluated in a quasi-analytical method. We generate 100 IEEE CM3 channel realizations, and for each channel realization, we use Table 2.1 to evaluate its $P_{D,1}$ performance and average the obtained $P_{D,1}$'s. In the experimental way, we still employ 100 IEEE CM3 channel realizations. For each realization, we generate 1000 test statistics to compare with the threshold and count the probability of detection. In order to evaluate the detection performance, we divide the SNR into three ranges. For example, when $P_{FA} = 0.1$, the low SNR range is below 0dB, the medium range is from 0dB to 6dB and the high SNR range is above 6dB. According to Fig. 2.7, the $P_{D,1}$ of $T(\mathbf{x})$ with $P = 3$ (solid line with $*$ markers) and the $P_{D,1}$ of $T_1'(\mathbf{x})$ (dash-dotted line with $*$ markers) are similar in the low and high SNR range. But in the medium range, $T(\mathbf{x})$ with $P = 3$ outperforms $T_1'(\mathbf{x})$ for about $5\% \sim 10\%$. For $P_{FA} = 10^{-3}$ and $P_{FA} = 10^{-5}$, the performance differences for these test statistics are large in the SNR range from 2dB to 8dB. $T(\mathbf{x})$ (solid lines with $\circ$ or $\diamond$ markers) can have a detection probability as high as 20% more than $T_1'(\mathbf{x})$ (dash-dotted lines with $\circ$ or $\diamond$ markers) under $\mathcal{H}_1$. However, when the test
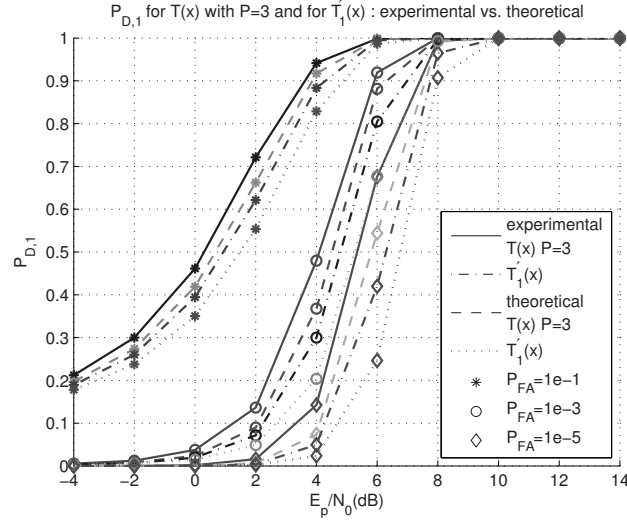
Figure 2.7: Experimental and theoretical $P_{D,1}$ performance comparison for $T(\mathbf{x})$ with $P = 3$ and $T_1'(\mathbf{x})$

statistic $T(\mathbf{x})$ is employed, we have to estimate the channel energy profile first. On the other hand, if we use the test statistic $T_1'(\mathbf{x})$, we only have to sum up the samples, which is easy to implement. But these results are only the detection probabilities under $\mathcal{H}_1$, which are used as boundaries for the overall performance under real circumstances.

As we have mentioned before, $P_{D,1}$ and $P_{D,1} + (1 - P_{D,1})P_{D,1}$ can be used as a lower boundary and an upper boundary for the overall P_{D_o}, respectively. We run Monte Carlo simulations to evaluate the P_{D_o} under real circumstances. For each run, the timing offset is randomly generated following a uniform distribution in the range of M_1 symbols, meanwhile the channel realization remains the same in order to exclude the channel influence in the multi-hypotheses case. In the detection procedure, once the first detection fails, we jump into the next observation window. When the second detection fails again, we declare a missed detection. The simulation results are shown in Fig. 2.8. The P_{D_o}'s of $T(\mathbf{x})$ with $P = 3$ (solid lines) lie between two boundaries: the upper boundaries (dashed lines) and the lower boundaries (dotted lines), and these boundaries are getting tighter as the P_{FA}'s are getting smaller. The P_{D_o}'s of $T_1'(\mathbf{x})$ (dash-dotted lines) are a bit higher than the P_{D_o}'s of $T(\mathbf{x})$. Especially for $P_{FA} = 10^{-3}$, around SNR= 6dB, the P_{D_o} of $T_1'(\mathbf{x})$ (dash-dotted line with ∘ markers) is 5% larger than the P_{D_o} of $T(\mathbf{x})$ (solid line with ∘ markers). That is because $T(\mathbf{x})$ weights each sample only based on two hypotheses $\mathcal{H}_0$ and $\mathcal{H}_1$. The weighting coefficients are not optimal for other hypotheses. The noise samples may be mistakenly weighted
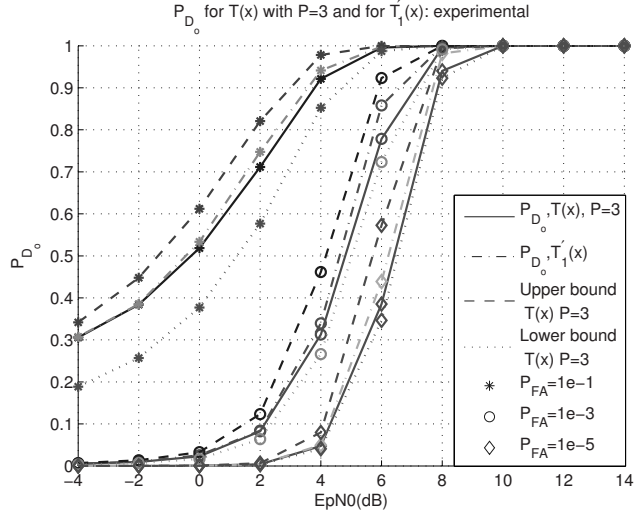
Figure 2.8: Experimental P_{D_o} for $T(\mathbf{x})$ with $P = 3$ and $T_1'(\mathbf{x})$

heavily under real circumstances. On the other hand, $T_1'(\mathbf{x})$ accumulates all the frame samples in the observation window, which is equivalent to equally weighting. According to these results, we can employ $T_1'(\mathbf{x})$ because of its simplicity and similar performance as $T(\mathbf{x})$.

500 Monte Carlo runs are used to evaluate the mean squared error (MSE) of $\hat{\mathbf{h}}_\delta$ vs. SNR. In each run, the timing offset and the channel are randomly generated. The results for the symbol-long estimates and the L_w-long estimates assuming perfect timing are shown in Fig. 2.9. The MF curves (dotted lines) always have the highest noise floor, since the MF output is the convolution of the channel energy vector with the code autocorrelation function. The performance gap for symbol-long estimates between the LS/LMMSE (dashed lines/solid lines) estimator and the MF is large. When we concentrate on the channel estimates in a limited range, such as 30 ns (lines with ○ markers) and 90 ns (lines with ◇ markers), the gap between the MF and the LS/LMMSE estimator is smaller. The normalized MSE $E[|(\hat{\delta} - \delta)/L_s|^2]$ for δ estimation is also assessed with different values of L_w based on different channel estimators. From Fig. 2.10, we see that the δ estimates based on MF (dotted lines), LS (dashed lines) and LMMSE (solid lines) channel estimates with the same L_w have similar performance and $L_w = 30$ ns is the best choice among all. The MSE for δ with L_w=30 ns (lines with ○ markers) is saturated after the SNR reaches 10dB. This is because we use NLOS channels, where the first path may not be the strongest and there is always remaining a fractional timing offset ϵ. Meanwhile the differences of the MSE for channel estimation with a 90ns range based on different methods (lines with ◇
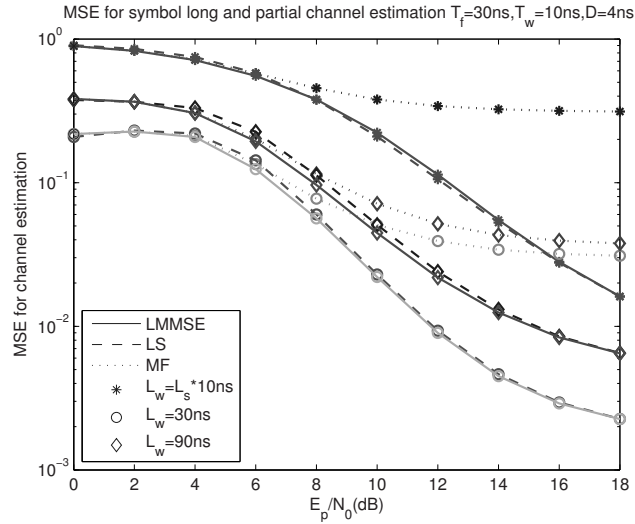
Figure 2.9: MSE Performance for channel estimation with different lengths

markers) are quite small around 10dB in Fig. 2.9, which will be employed to construct the equalizer. As a result, we choose the MF as the channel estimator.

Furthermore, combinations of the MF channel estimator with different equalizers are investigated. We employ $L_w = 30$ ns for synchronization. Fig. 2.11 shows the BER performance. The BER performance for the MF equalizer (lines with ∘ markers) approaches 0 after 12dB, while the performances for the ZF (lines with ∗ markers) and the LMMSE equalizers (lines with □ markers) approach 0 after 10dB. Hence, the MF equalizer is 2dB worse than the ZF and the LMMSE equalizer, and all of them employ 90 ns long channel estimates. The curves of the ZF equalizer and the LMMSE equalizer overlay each other. The bias does not have much impact on them. They have almost the same performance. As a result, the optimal combination considering cost and performance would be a MF channel estimator with a ZF equalizer. According to the results above, we can remark that the IFI after the I& D is not so serious in our simulation setup, since the channel energy attenuates exponentially and one frame contains most of the energy. The performance differences of different equalizers are not so obvious. However, the LMMSE estimator has the potential to handle more serious IFI and ISI. The effects of the bias on the BER performance can be ignored, but it has to be taken into account for the channel estimation (done implicitly, see Section 2.4.1). When we want to shorten the frame length to achieve a higher data rate, more interference will be generated. We then need a more accurate data model to handle this interference.
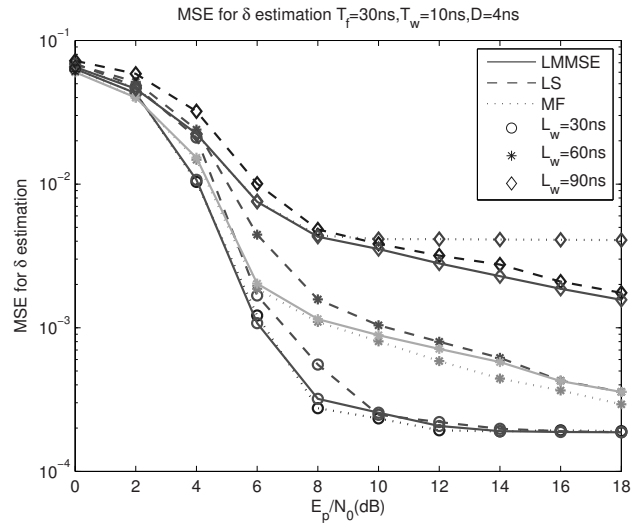
Figure 2.10: MSE Performance for δ estimation with various L_w's
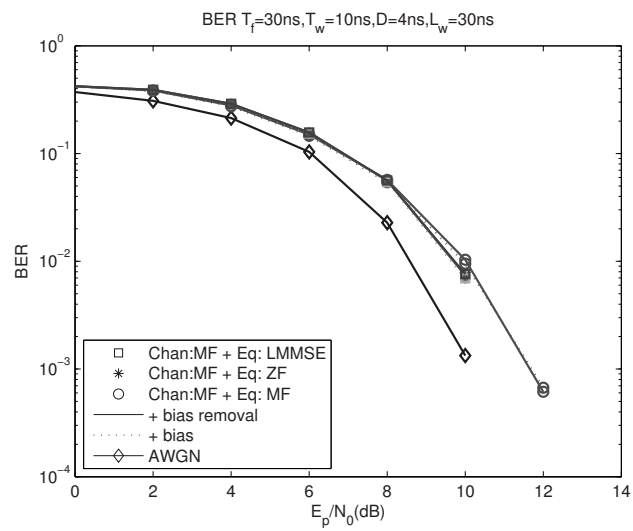
Figure 2.11: BER Performance for CM3

2.6   Conclusions

We have proposed a complete solution for signal detection, channel estimation, synchronization and equalization in a TR-UWB system. The scheme is based on a data model, which takes IPI, IFI and ISI into account and releases the frame time requirements to allow for higher data rate communications. Several detectors based on a specific training scheme are derived and assessed. We find that the simple detector, which sums up all the samples in the observation window and compares the result with a threshold, gives a good balance between performance and cost. Moreover, the joint channel and timing estimation is achieved in three different ways. The property of the circulant matrix in the data model is exploited to reduce the complexity of the algorithms. Then a two-stage synchronization strategy is proposed to first achieve sample-level synchronization and later symbol-level synchronization. Last but not least, three kinds of equalizers are derived. We evaluate different combinations of channel estimation and equalization schemes using the IEEE UWB channel model CM3, which shows that the TR-UWB system can be implemented with low cost and achieve moderate data rate communications.

2.A   Noise analysis

The noise autocorrelation term $n_0[n]$ is

$$n_0[n] = \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} n(t)n(t+D)dt, \qquad (2.58)$$

where $n(t)$ is band limited AWGN and its autocorrelation function is $R_n(\tau) = E[n(t)n(t-\tau)] = N_0 B \mathrm{sinc}(2B\tau)$. Therefore, $n_0[n]$ has approximately zero mean, as a result of $R_n(D) \approx 0$ based on the assumption $D \gg 1/B$. According to the Gaussian joint variable theorem [59] [60], its variance can be derived as

$$\begin{aligned}
\mathrm{var}(n_0[n]) &\approx E\left[n_0^2[n]\right] \\
&\approx \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} \left[R_n^2(t-u) + R_n(t-u-D)R_n(t+D-u)\right] dt\, du.
\end{aligned}$$
$$(2.59)$$

The second term is the product of two sinc functions offset by $2D$, which is approximately zero by using the property of sinc function saying that $\mathrm{sinc}(2B\tau)\mathrm{sinc}(2B(\tau+\Delta)) \approx \mathrm{sinc}^2(2B\tau)\delta(\Delta)$, where $\delta(\Delta)$ is the Kronecker delta. Recalling $R_n(D) \approx 0$ and $T_{\mathrm{sam}} \gg 1/B$, and applying Parseval's theorem, we derive the variance of $n_0[n]$ as (also see [61])

$$\begin{aligned}
\mathrm{var}(n_0[n]) &\approx \frac{N_0^2}{4} \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} \left[4B^2\mathrm{sinc}^2(2B(t-u))\right] dt\, du \\
&\approx \frac{N_0^2}{4} \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} \left[\int_{-B}^{B} 1 df\right] dt \\
&= \frac{N_0^2 B T_{\mathrm{sam}}}{2}.
\end{aligned}$$
$$(2.60)$$

In summary, $n_0[n]$ is approximately zero mean and white with variance $N_0^2 B T_{\mathrm{sam}}/2$. These noise autocorrelation samples are uncorrelated with each other, due to the assumption $T_{\mathrm{sam}} \gg 1/B$.

Furthermore, the aggregate noise term $n_1[n]$ is

$$\begin{aligned}
n_1[n] &= n_0[n] + s_i c_j \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} [h(t-\tau)n(t) + h(t-D-\tau)n(t+D)]dt \\
&\quad + \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} [h(t-\tau)n(t+D) + h(t+D-\tau)n(t)]dt.
\end{aligned}$$
$$(2.61)$$

Defining

$$\gamma'[n] = s_i c_j \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h(t-\tau)n(t) + h(t-D-\tau)n(t+D)]dt, \quad (2.62)$$

$$\gamma''[n] = \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h(t-\tau)n(t+D) + h(t+D-\tau)n(t)]dt, \quad (2.63)$$

we obtain

$$n_1[n] = \gamma'[n] + \gamma''[n] + n_0[n]. \quad (2.64)$$

$\gamma'[n]$ and $\gamma''[n]$ are random variables, resulting from the cross-correlation between the signal and the noise.

Now we will derive the statistical properties of these two random variables. Both $\gamma'[n]$ and $\gamma''[n]$ have zero mean. The variance of $\gamma'[n]$ is calculated as follows

$$\text{var}(\gamma'[n]) = E[|\gamma'[n]|^2] \quad (2.65)$$

$$= \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} [h(t-\tau)h(u-\tau)R_n(t-u)$$

$$+ h(t-D-\tau)h(u-D-\tau)R_n(t-u)]dtdu. \quad (2.66)$$

Let us insert $R_n(\tau)$ into the first term (also see [61])

$$\int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} h(t-\tau)h(u-\tau)R_n(t-u)dtdu$$

$$= \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} h(t-\tau)h(u-\tau)N_0 B\text{sinc}(2B(t-u))dtdu$$

$$= \frac{N_0}{2} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} h(t-\tau)h(u-\tau) \int_{-B}^{B} e^{j2\pi f(t-u)}df dt du$$

$$= \frac{N_0}{2} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} h(t-\tau) \int_{-B}^{B} e^{j2\pi f(t-\tau)}df dt$$

$$\times \int_{(n-1)T_{\text{sam}}-\tau}^{nT_{\text{sam}}-\tau} h(u-\tau)e^{-j2\pi f(u-\tau)}d(u-\tau)$$

$$= \frac{N_0}{2} \int_{(n-1)T_{\text{sam}}}^{nT_{\text{sam}}} h(t-\tau) \left(\int_{-B}^{B} H(f)e^{j2\pi f(t-\tau)}df \right) dt, \quad (2.67)$$

where $H(f)$ is the Fourier transform of $h(u-\tau), u \in [(n-1)T_{\text{sam}}, nT_{\text{sam}}]$, which is a segment of the aggregate channel. Since the bandwidth B of $n(t)$ is assumed much larger than the bandwidth of $h(u-\tau), u \in [(n-1)T_{\text{sam}}, nT_{\text{sam}}]$, we obtain $\int_{-B}^{B} H(f)e^{j2\pi f(t-\tau)}df \approx h(t-\tau), t \in [(n-1)T_{\text{sam}}, nT_{\text{sam}}]$. As a result, we obtain

similar results as in [55] [56] [61]

$$\int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} h(t-\tau)h(u-\tau)R_n(t-u)dtdu$$

$$\approx \frac{N_0}{2} \int_{(n-1)T_{\mathrm{sam}}}^{nT_{\mathrm{sam}}} h(t-\tau)h(t-\tau)dt$$

$$= \frac{N_0}{2} R(0, n-\delta). \tag{2.68}$$

In a similar way, the other term of $\mathrm{var}(\gamma'[n])$ can be deduced. The same method is applied to $\mathrm{var}(\gamma''[n])$ and $E\big[\gamma'[n]\gamma''[n]\big]$. All the derivations are based on the assumption that $R_n(D) \approx 0$ and $T_{\mathrm{sam}} \gg 1/B$. The results are summarized as follows

$$\mathrm{var}(\gamma'[n]) \approx \begin{cases} \frac{N_0}{2}\left(R(0, n-\delta) + R(0, n-\delta-\frac{D}{T_{\mathrm{sam}}})\right), \\ \qquad n = \delta+1, \dots, \delta + P_h, \\ 0, \quad \text{elsewhere}, \end{cases} \tag{2.69}$$

$$\mathrm{var}(\gamma''[n]) = E[|\gamma''[n]|^2]$$

$$\approx \begin{cases} \frac{N_0}{2}\left(R(0, n-\delta) + R(0, n-\delta+\frac{D}{T_{\mathrm{sam}}})\right), \\ \qquad n = \delta+1, \dots, \delta + P_h, \\ 0, \quad \text{elsewhere}, \end{cases} \tag{2.70}$$

$$E\big[\gamma'[n]\gamma''[n]\big] \approx \begin{cases} \frac{N_0}{2}s_i c_j\left(R(D, n-\delta) + R(D, n-\delta+\frac{D}{T_{\mathrm{sam}}})\right), \\ \qquad n = \delta+1, \dots, \delta + P_h, \\ 0, \quad \text{elsewhere}, \end{cases} \tag{2.71}$$

$$E\big[\gamma'[n]n_0[n]\big] = E\big[\gamma''[n]n_0[n]\big] = 0. \tag{2.72}$$

In summary, the stochastic properties of $n_1[n]$ are

$$E\big[n_1[n]\big] \approx 0, \tag{2.73}$$

$$\mathrm{var}(n_1[n]) \approx \begin{cases} \frac{N_0}{2}\Big\{2R(0, n-\delta) + R(0, n-\delta-\frac{D}{T_{\mathrm{sam}}}) + R(0, n-\delta+\frac{D}{T_{\mathrm{sam}}}) \\ \quad + s_i c_j\left(2R(D, n-\delta) + 2R(D, n-\delta+\frac{D}{T_{\mathrm{sam}}})\right)\Big\} + \sigma_0^2, \\ \qquad n = \delta+1, \dots, \delta + P_h, \\ 0, \quad \text{elsewhere} \end{cases} \tag{2.74}$$

where $\sigma_0^2 = N_0^2 B T_{\text{sam}}/2$. These aggregate noise samples are uncorrelated with each other, recalling that $T_{\text{sam}} \gg 1/B$. This assumption has usually been satisfied by UWB signals (e.g., in our case $T_{\text{sam}} = 10\text{ns}$, $B \approx 2/T_p = 10\text{GHz}$, then $2BT_{\text{sam}} = 200$). Also $n_0[n]$ and $n_1[n]$ can be assumed as Gaussian random variables by invoking the sampling theorem and the central limit theorem [59].

2.B Detector derivation

In summary, the statistics of $\mathbf{x}$ in (2.31) are

$$\mathcal{H}_0 : \mathbf{x} \overset{a}{\sim} \mathcal{N}(0, \sigma_0^2 \mathbf{I}), \tag{2.75}$$

$$\mathcal{H}_1 : \mathbf{x} \overset{a}{\sim} \mathcal{N}(\mathbf{1}_{M_1 N_f} \otimes \mathbf{z}_f, \text{diag}(\boldsymbol{\lambda})). \tag{2.76}$$

The *Neyman-Pearson* detector decides $\mathcal{H}_1$ if

$$L(\mathbf{x}) = \frac{p(\mathbf{x}; \mathcal{H}_1)}{p(\mathbf{x}; \mathcal{H}_0)} > \gamma, \tag{2.77}$$

where γ is found by making the probability of false alarm P_{FA} to satisfy

$$P_{FA} = Pr\{L(\mathbf{x}) > \gamma; \mathcal{H}_0\} = \alpha. \tag{2.78}$$

$L(\mathbf{x})$ can be expressed as

$$L(\mathbf{x}) = \frac{\displaystyle\prod_{i=1}^{P} \frac{\exp\left[-\dfrac{1}{2(2N_0 z_f[i] + \sigma_0^2)} \displaystyle\sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} (x[nP+i] - z_f[i])^2\right]}{(2\pi(2N_0 z_f[i] + \sigma_0^2))^{\frac{M_1 N_f}{2}}}}{\dfrac{1}{(2\pi\sigma_0^2)^{\frac{M_1 N_f P}{2}}} \exp\left[-\dfrac{1}{2\sigma_0^2} \displaystyle\sum_{n=(k-1)N_f P+1}^{(k+M_1-1)N_f P} x^2[n]\right]}. \tag{2.79}$$

Defining $\sigma_1^2[i] = 2N_0 z_f[i] + \sigma_0^2$, inserting it into $\ln L(\mathbf{x})$ and eliminating the con-

stants leads to

$$
\ln L(\mathbf{x})
$$

$$
= \sum_{i=1}^{P} \left\{ \frac{1}{2\sigma_0^2} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i] - \frac{1}{2\sigma_1^2[i]} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} (x[nP+i] - z_f[i])^2 \right\}
$$

$$
= \sum_{i=1}^{P} \left\{ \frac{2z_f[i]}{2\sigma_1^2[i]} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x[nP+i] + \left(\frac{1}{2\sigma_0^2} - \frac{1}{2\sigma_1^2[i]} \right) \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i] \right\}
$$

$$
= \sum_{i=1}^{P} \left\{ \frac{z_f[i]}{\sigma_1^2[i]} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x[nP+i] + \frac{N_0 z_f[i]}{\sigma_0^2 \sigma_1^2[i]} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i] \right\}
$$

$$
= \sum_{i=1}^{P} \frac{z_f[i]}{\sigma_1^2[i]} \left\{ \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x[nP+i] + \frac{N_0}{\sigma_0^2} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i] \right\}.
$$

$$(2.80)$$

Then the test statistic is

$$
T(\mathbf{x}) = \sum_{i=1}^{P} \frac{z_f[i]}{\sigma_1^2[i]} \left\{ \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x[nP+i] + \frac{N_0}{\sigma_0^2} \sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i] \right\}.
$$

$$(2.81)$$

2.C Statistic of the detectors

2.C.1 Detector $T_1(\mathbf{x})$

Since $\mathbf{x}$ is assumed to be a Gaussian vector, $T_1(\mathbf{x})$ also follows a Gaussian distribution.

$$
\mathcal{H}_0 : T_1(\mathbf{x}) \overset{a}{\sim} \mathcal{N}(0, M_1 N_f \sigma_0^2 \sum_{i=1}^{P} \frac{z_f^2[i]}{\sigma_1^4[i]}), \tag{2.82}
$$

$$
\mathcal{H}_1 : T_1(\mathbf{x}) \overset{a}{\sim} \mathcal{N}(M_1 N_f \sum_{i=1}^{P} \frac{z_f^2[i]}{\sigma_1^2[i]}, M_1 N_f \sum_{i=1}^{P} \frac{z_f^2[i]}{\sigma_1^2[i]}). \tag{2.83}
$$

Actually, if the condition $z_f[i]/N_0 \ll BT_{\text{sam}}/4$ is satisfied, which means the signal-to-noise ratio (SNR) is low, the term $2N_0 z_f[i]$ can be ignored in the variance of $\mathbf{x}$ under $\mathcal{H}_1$, and then $T_1(\mathbf{x})$ can be derived directly.

2.C.2   Detector $T_2(\mathbf{x})$

Since the different entries of $\mathbf{x}$ have different weighting factors in $T_2(\mathbf{x})$, we collect the data samples bearing the same weighting factor into the same group. Therefore, there are P groups of data samples and they are assumed to be uncorrelated. Each group $\sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} x^2[nP+i]$ follows a chi-squared distribution. However, $T_2(\mathbf{x})$ is still assumed to be a Gaussian variable, as it is the sum of the weighted groups. Then we can obtain

$$
\mathcal{H}_0:
$$
$$
\sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} \frac{x^2[nP+i]}{\sigma_0^2} \overset{a}{\sim} \chi^2_{M_1 N_f},
$$
$$
T_2(\mathbf{x}) \overset{a}{\sim} \mathcal{N}\left(M_1 N_f \sigma_0{}^2 \sum_{i=1}^{P} \frac{z_f[i]}{\sigma_1^2[i]}, \; 2M_1 N_f \sigma_0{}^4 \sum_{i=1}^{P} \frac{z_f^2[i]}{\sigma_1^4[i]}\right),
$$
$$
\mathcal{H}_1:
$$
$$
\sum_{n=(k-1)N_f}^{(k+M_1-1)N_f-1} \frac{x^2[nP+i]}{\sigma_1^2[i]} \overset{a}{\sim} \chi^2_{M_1 N_f}\left(\frac{M_1 N_f E_f^2[i]}{\sigma_1^2[i]}\right),
$$
$$
T_2(\mathbf{x}) \overset{a}{\sim} \mathcal{N}\left(M_1 N_f \sum_{i=1}^{P} z_f[i]\left(1+\frac{z_f^2[i]}{\sigma_1^2[i]}\right),\right.
$$
$$
\left. 2M_1 N_f \sum_{i=1}^{P} z_f^2[i]\left(1+\frac{2z_f^2[i]}{\sigma_1^2[i]}\right)\right),
$$

where χ^2_ν is the central chi-squared pdf with ν degrees of freedom which has mean ν and variance 2ν. Meanwhile, $\chi^2_\nu(\lambda)$ is the non-central chi-squared pdf with ν degrees of freedom and noncentrality parameter λ. Hence, it has mean $\nu + \lambda$ and variance $2\nu + 4\lambda$.

Part II

UWB Ranging
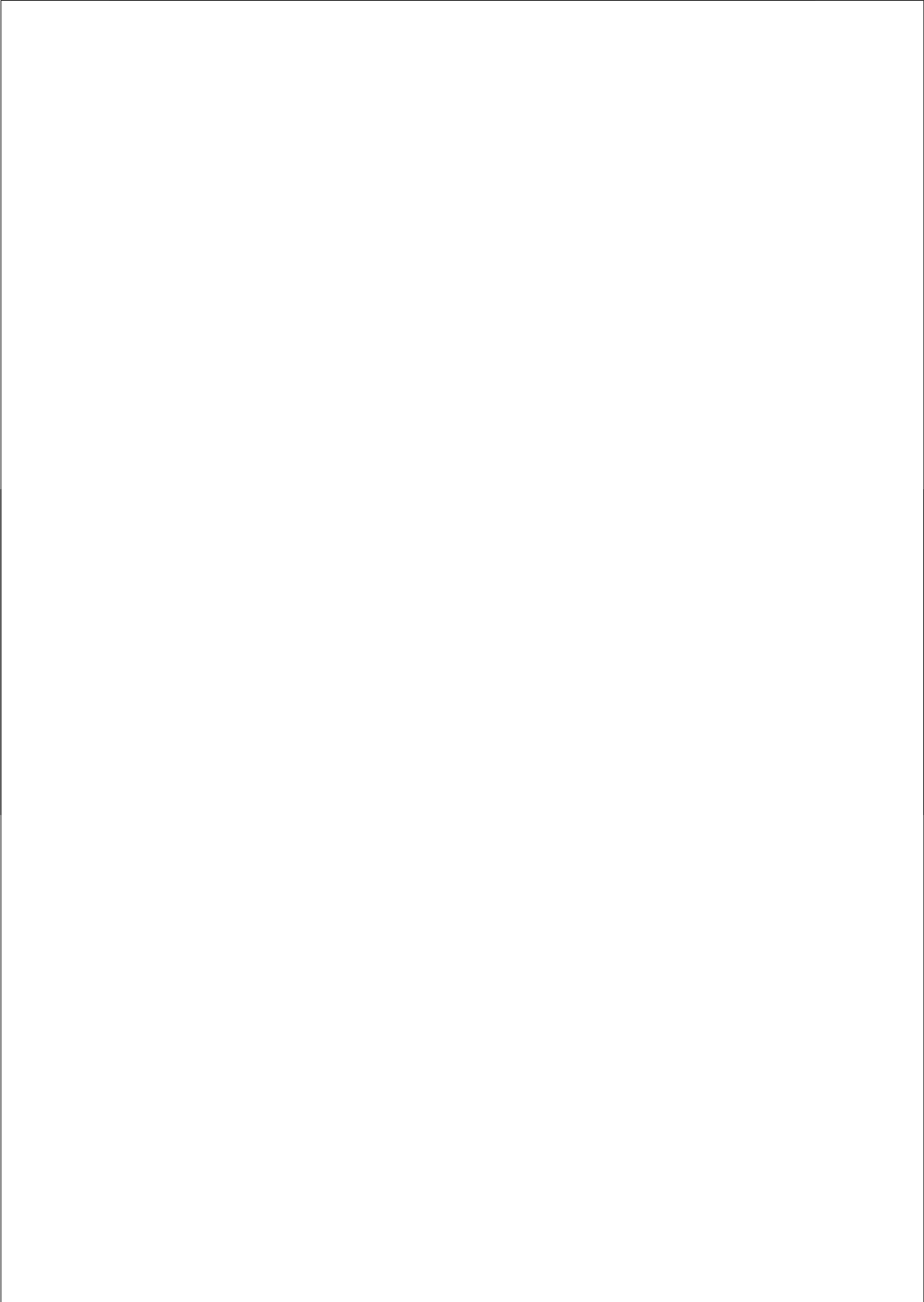
Cramér-Rao Bound for Range Estimation

3.1 Introduction

The Cramér-Rao bound (CRB) is a standard benchmark to evaluate the performance of an estimator. In this chapter, we investigate the CRB for range (D) estimation (the distance between the transmitter and the receiver), which is an important parameter for localization. There are two elementary measurements for range estimation: the received signal strength (RSS) and the time of arrival (TOA–τ) [1]. The existing literature mostly treats them separately to derive the CRB(D), e.g., see [1] [63]. Some of them [64] insert the path-loss model (indicating the relationship between the RSS and the range) into the received signal-to-noise ratio (SNR), which is viewed as a parameter of the CRB(τ)(in this case, we have CRB(D) = c^2CRB(τ)). However, they do not exploit the range information in the RSS and the TOA jointly. Previous work in [65] proposes to use both the RSS and the TOA to improve the ranging accuracy and derives a CRB(D) by simply, yet incorrectly, assuming that they are uncorrelated. Moreover, it does not provide a method to combine them. Other joint methods [66] are for location estimation. They propose to fuse the TOA measurements and the RSS measurements to get a lower bound for localization, meanwhile they assume the estimation error of the TOA has constant variance regardless of its distance dependency.

We investigate the relationship among the RSS, the TOA and the range D, and use both the RSS and the TOA for the CRB(D). We do not use the RSS explicitly as a parameter, but the amplitude of the received signal. Single path propagation is assumed, yet the path-loss model is taken into account in the

---

The results in this chapter appeared in [62].

received signal model to explore the distance dependency of the amplitude of the received signal. The amplitude and the TOA are represented as a function of D explicitly in the received signal model. Thus, we can derive the CRB(D) directly.

The rest of the chapter is organized as follows. In Section 3.2, we derive the true CRB(D) in an additive white Gaussian noise (AWGN) channel with attenuation. Some results are shown. Further, the maximum likelihood estimator (MLE) of D is proposed. In Section 3.3, we include shadowing. Due to the difficulty to derive the true CRB(D) directly in this case, we propose several CRB-like bounds and suboptimal estimators for D. Numerical results are also shown in Section 3.3. We conclude the chapter in Section 3.4.

3.2   CRB in an AWGN channel with attenuation

The received signal through an AWGN channel with an attenuation coefficient is

$$r(t) = \alpha s(t - \tau) + n(t), \tag{3.1}$$

where τ is the unknown deterministic time delay, which is related to the range D as $\tau = D/c$, with c the signal propagation speed. The channel attenuation coefficient α is related to D following the distance-power law [67] as $\alpha = k_0/\sqrt{D^{\gamma_0}}$, with k_0 a constant parameter and γ_0 also a constant depending on the environment, e.g., in free space $\gamma_0 = 2$. The transmitted waveform is represented by $s(t)$. We use $n(t)$ to denote AWGN with double-sided power spectral density $N_0/2$, which is filtered by an ideal anti-aliasing filter of bandwidth B. We assume B is larger than the signal bandwidth. The continuous signal is sampled at the Nyquist rate $2B = 1/T_s$. The received data samples are collected in a vector $\mathbf{r} = [r(0), r(T_s), r(2T_s), \ldots, r((N-1)T_s)]^T$, which can be written as

$$\mathbf{r} = k_0 D^{-\frac{\gamma_0}{2}} \mathbf{s}_D + \mathbf{n}, \tag{3.2}$$

where $\mathbf{s}_D = [s(-\tau), s(T_s - \tau), \ldots, s((N-1)T_s - \tau)]^T$ and $\mathbf{n}$ contains noise samples with variance $\sigma^2 = N_0 B$. In this model, we find the range information not only in the time delay, but also in the amplitude of the received signal. Therefore, the CRB(D) based on the data model (3.2) is

$$\mathrm{CRB}(D) = \left\{ E_{\mathbf{r}} \left[-\frac{\partial^2}{\partial D^2} \ln(p(\mathbf{r}; D)) \right] \right\}^{-1}. \tag{3.3}$$

with $p(\mathbf{r}; D)$ following a Gaussian distribution [68]. Assuming the observation window includes the whole waveform, it leads to (see Appendix 3.A for derivation

details)

$$\text{CRB}(D) = \frac{c^2 D^{\gamma_0+2}}{2k_0^2 \frac{\mathcal{E}}{N_0} \left(\frac{\gamma_0^2 c^2}{4} + D^2 \overline{F^2} + \gamma_0 c D \overline{F'^2} \right)}. \tag{3.4}$$

where T_o is the length of the observation window ($T_o = NT_s$), $\mathcal{E} = \int_0^{T_o} s^2(t)dt$ is the transmitted signal energy, $\overline{F^2} = \int_{-\infty}^{\infty} (2\pi F)^2 |S(F)|^2 dF \Big/ \int_{-\infty}^{\infty} |S(F)|^2 dF$ is the mean square bandwidth of the signal, with $S(F)$ the Fourier transform of $s(t)$, and $\overline{F'^2} = \int_0^{T_o} s(t) \frac{ds(t)}{dt} dt \Big/ \int_0^{T_o} s^2(t)dt = \frac{s^2(t)}{2} \Big|_0^{T_o} \Big/ \mathcal{E}$. Since $s(t)$ is smooth in $[0, T_o)$, we may assume $s(0) = s(T_o)$. As a result, $\overline{F'^2} = 0$. Therefore, the CRB(D) in (3.4) can be simplified as

$$\text{CRB}(D) = \frac{c^2 D^{\gamma_0+2}}{2k_0^2 \frac{\mathcal{E}}{N_0} \left(\frac{\gamma_0^2 c^2}{4} + D^2 \overline{F^2} \right)}. \tag{3.5}$$

We now compare CRB(D) in (3.5) with the results in [64]. We first apply the method in [64] to derive the CRB(τ), and then obtain the CRB(D) as a transformation of the CRB(τ). It results into

$$\text{CRB}_{ref}(D) = \frac{c^2 D^{\gamma_0}}{2k_0^2 \frac{\mathcal{E}}{N_0} \overline{F^2}}. \tag{3.6}$$

Though the method in [64] takes the path-loss model into account, it only exploits the range information in the time delay. The CRB(D) in (3.5) has one more positive term $\gamma_0^2 c^2/4$ in the denominator, as a result of additionally investigating the range information in the amplitude of the received signal. Setting $\gamma_0^2 c^2/4 = D^2 \overline{F^2}$, we can obtain a critical distance $D_c = \gamma_0 c/(2\sqrt{\overline{F^2}})$, which is determined by the environment, the signal propagation speed and the mean square bandwidth of the transmitted signal. When $D \gg D_c$, $\text{CRB}(D) \approx \text{CRB}_{ref}(D)$. On the other hand, when $D \approx D_c$, we can gain about 3dB by taking the additional information into account.

Now, we will give some examples. It is known that $\mathcal{E} = \int_0^{T_o} s^2(t)dt = T_o \int_{-\infty}^{\infty} \Phi(F)dF$, where $\Phi(F)$ is the power spectral density of the signal. Assuming $\Phi(F)$ is uniformly distributed over the bandwidth ($W = f_H - f_L$) of the signal, we get $\overline{F^2} = 4\pi^2(\frac{W^2}{3} + f_H f_L)$. A larger bandwidth and a higher central frequency (or a higher carrier frequency) result in a larger $\overline{F^2}$ and a lower CRB. Using $k_0 = 1$, $\gamma_0 = 2$, $c = 3 \cdot 10^8$m/s and the whole bandwidth ($W = 10.6$GHz $- 3.1$GHz) of the ultra-wide band (UWB) signals, D_c is approximately 6.6mm, which is quite small. Hence, for an indoor environment, where D is in the range of a few meters,

D(m)	UWB signals σ_D(m)			D(m)	NB signals, 0MHz $\sim$ 5MHz	
	3.1GHz $\sim$ 10.6GHz	7.316GHz $\sim$ 8.684GHz	3.658GHz $\sim$ 4.342GHz		σ_D(m)	σ_{Dref}(m)
1	$9.0786 \cdot 10^{-8}$	$1.9066 \cdot 10^{-7}$	$5.3924 \cdot 10^{-7}$	10	$1.4317 \cdot 10^{-2}$	$2.7671 \cdot 10^{-2}$
3	$2.7236 \cdot 10^{-7}$	$5.7199 \cdot 10^{-7}$	$1.6178 \cdot 10^{-6}$	15	$2.7884 \cdot 10^{-2}$	$4.1507 \cdot 10^{-2}$
10	$9.0788 \cdot 10^{-7}$	$1.9066 \cdot 10^{-6}$	$5.3927 \cdot 10^{-6}$	20	$4.2648 \cdot 10^{-2}$	$5.5343 \cdot 10^{-2}$

Table 3.1: Theoretical ranging accuracy for an AWGN channel with attenuation, $\sigma_D = \sqrt{\mathrm{CRB}(D)}$

it is sufficient to consider the range information in the delay only. This is due to the large bandwidth and the high central frequency of the UWB signals. However, if we only use a narrowband (NB) signal, for example, with a bandwidth of 5MHz and a carrier frequency of 2.5MHz, D_c would be 16.5m under those circumstances. Hence, for an indoor environment, we can benefit from taking the range information in both the amplitude and the delay into account.

The left half of Table 3.1 shows the CRB(D) (3.5) in an AWGN channel with attenuation employing UWB signals with different bandwidths. The transmitted power spectral density of the signal $\Phi(F)$ is restricted below the FCC mask (-41.3 dBm/MHz). Further, $k_0 = 1, \gamma_0 = 2, c = 3 \cdot 10^8$m/s, $N_0 = 2 \cdot 10^{-20}$W/Hz, and $T_o = 100$ns. Clearly, increasing the bandwidth and central frequency, we obtain a more accurate range estimation. Meanwhile, the right half of Table 3.1 compares the new CRB(D) (3.5) with the conventional CRB$_{ref}$(D) (3.6) employing NB signals. The NB signals have a frequency range from 0 to 5MHz (a bandwidth of 5MHz and a carrier frequency of 2.5MHz). The observation time T_o is 1μs. Other parameters are kept the same. We can see the new CRB(D) (3.5) is almost 3dB lower than the conventional CRB$_{ref}$(D)   (3.6).

Let us now derive the MLE of D, which can asymptotically attain the bound in (3.5). We would like to find the D that maximizes $p(\mathbf{r}; D)$ or $\ln p(\mathbf{r}; D)$, leading to

$$\hat{D} = \operatorname*{argmin}_{0 < D < cT_o} \left\{ k_0 D^{-\gamma_0} \mathcal{E} - 2D^{-\frac{\gamma_0}{2}} \mathcal{E}_{rs}(D) \right\}. \tag{3.7}$$

where $\mathcal{E}_{rs}(D)$ is a function of D: $\mathcal{E}_{rs}(D) = \int_0^{T_o} r(t)s(t - D/c)dt$. A grid search is then executed to look for D. The variance of this estimator approaches (3.5) as long as the data record is large enough.

3.3   CRBs in an AWGN channel with attenuation and shadowing

When we include shadowing effects, which represent a more realistic environment, the received signal is

$$r(t) \quad = \quad \widetilde{\alpha} X s(t - \tau) + n(t). \tag{3.8}$$

where X is a random variable modeling the shadowing effects and following a lognormal distribution $20\log_{10}X \sim \mathcal{N}(0, \sigma_x^2)$, and $\widetilde{\alpha}$ still follows a distance-power law as $\widetilde{\alpha} = \widetilde{k_0}^2 / \sqrt{D^{\gamma_0}}$. In order to get a fair comparison, we normalize the average channel energy, resulting in $\widetilde{\alpha}^2 E[X^2] = \alpha^2$. Therefore, $\widetilde{k_0}^2 = k_0^2 / E[X^2]$. The normalization excludes the influence of the shadowing coefficient to the average received energy. The definitions for other parameters remain the same as in the last section. The received signal vector is

$$\mathbf{r} \quad = \quad \widetilde{k}_0 D^{-\frac{\gamma_0}{2}} X \mathbf{s}_D + \mathbf{n}. \tag{3.9}$$

Since X is independent of D, it can be viewed as a nuisance parameter. The MLE for D tries to find the D that maximizes $p(\mathbf{r}; D)$. It is known that $p(\mathbf{r}; D) = \int p(\mathbf{r}|X; D)p(X)dX$. By integration, we can get rid of X leading to a closed form of $p(\mathbf{r}; D)$, which is not only needed for the MLE, but also for the CRB(D). However, the integration is very difficult to derive in closed form.

Due to the difficulties to obtain the closed form of $p(\mathbf{r}; D)$, we will derive the CRB for $\boldsymbol{\theta} = [D, X]^T$. Since we have prior knowledge of X, the Bayesian information matrix (BIM) [68] for $\boldsymbol{\theta}$ is employed

$$\mathbf{I}_B(\boldsymbol{\theta}) \quad = \quad E_{\boldsymbol{\theta}}[\mathbf{I}_D(\boldsymbol{\theta})] + \mathbf{I}_P(\boldsymbol{\theta}), \tag{3.10}$$

$$[\mathbf{I}_D(\boldsymbol{\theta})]_{ij} \quad = \quad -E_{\mathbf{r}|\boldsymbol{\theta}} \left[\frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \ln p(\mathbf{r}|\boldsymbol{\theta}) \right], \tag{3.11}$$

$$[\mathbf{I}_P(\boldsymbol{\theta})]_{ij} \quad = \quad -E_{\boldsymbol{\theta}} \left[\frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \ln p(\boldsymbol{\theta}) \right]. \tag{3.12}$$

where $\mathbf{I}_D(\boldsymbol{\theta})$ represents information obtained from the data, $\mathbf{I}_P(\boldsymbol{\theta})$ indicates the prior information and $p(\mathbf{r}|\boldsymbol{\theta})$ follows a Gaussian distribution. Hence, we obtain

$$\mathbf{I}_D(\boldsymbol{\theta}) = \frac{2\widetilde{k_0}^2}{D^{\gamma_0}} \frac{\mathcal{E}}{N_0} \left[\begin{array}{cc} X^2 \left(\frac{\gamma_0^2}{4D^2} + \frac{\overline{F^2}}{c^2} + \frac{\gamma_0 \overline{F'^2}}{cD} \right) & -X \left(\frac{\gamma_0}{2D} + \frac{\overline{F'^2}}{c} \right) \\ -X \left(\frac{\gamma_0}{2D} + \frac{\overline{F'^2}}{c} \right) & 1 \end{array} \right]$$

$$\mathbf{I}_P(\boldsymbol{\theta}) = \left[\begin{array}{cc} 0 & 0 \\ 0 & X_c \end{array} \right], \quad \text{where } X_c = E_X \left[-\frac{\partial^2 \ln p(X)}{\partial X^2} \right].$$

$$
\begin{aligned}
\mathrm{HCRB}(D) \quad &= \quad A/B \tag{3.13}\\
A \quad &= \quad c^2 D^{\gamma_0+2}\\
B \quad &= \quad 2\widetilde{k_0}^{\,2}\frac{\mathcal{E}}{N_0}\left\{\left(M_2 - M_1^2 + \frac{M_1^2}{1+\dfrac{2\widetilde{k_0}^{\,2}}{X_c D^{\gamma_0}}\dfrac{\mathcal{E}}{N_0}}\right)\left(\frac{\gamma_0^2 c^2}{4}+\gamma_0 c D\overline{F'^2}\right)\right.\\
&\qquad\left. +M_2 D^2\overline{F^2} - M_1^2\left(1-\frac{1}{1+\dfrac{2\widetilde{k_0}^{\,2}}{X_c D^{\gamma_0}}\dfrac{\mathcal{E}}{N_0}}\right)D^2(\overline{F'^2})^2\right\},
\end{aligned}
$$

The derivation details are listed in Appendix 3.B.

We now have all the ingredients to derive several bounds. The first bound is the Hybrid CRB (HCRB) [69], which is $\mathrm{HCRB}(D) = \left[\mathbf{I}_B^{-1}(\boldsymbol{\theta})\right]_{11}$ as shown in (3.13), with $M_2 = E[X^2]$ and $M_1 = E[X]$. It covers the case where the desired deterministic parameter and random nuisance parameters are jointly estimated, and it is a bound for the estimators that take the prior knowledge of X into account.

The second bound is the Modified CRB (MCRB) [70]: $\mathrm{MCRB}(D)$ $= 1/[\mathbf{I}_B(\boldsymbol{\theta})]_{11}$, which is

$$
\mathrm{MCRB}(D) = \frac{c^2 D^{\gamma_0+2}}{E[X^2]2\widetilde{k_0}^{\,2}\dfrac{\mathcal{E}}{N_0}\left(\dfrac{\gamma_0^2 c^2}{4}+D^2\overline{F^2}+\gamma_0 c D\overline{F'^2}\right)}. \tag{3.14}
$$

It is a loose bound to cope with nuisance parameters, when it is difficult to get the true CRB. The MCRB(D) (3.14) depends on the average received energy. Taking $\widetilde{k_0}^{\,2} = k_0^2/E[X^2]$ into (3.14), we find it is equal to the CRB(D) (3.4). Due to the normalization, the average received energy in an AWGN channel with attenuation and shadowing is the same as the received energy in an AWGN channel only with attenuation.

The third bound is the Miller-Chang bound (MCB) proposed in [71]: $\mathrm{MCB}(D) = E_X\left\{1/[\mathbf{I}_D(\boldsymbol{\theta})]_{11}\right\}$, which is

$$
\mathrm{MCB}(D) = E_X\left[\frac{1}{X^2}\right]\frac{c^2 D^{\gamma_0+2}}{2\widetilde{k_0}^{\,2}\dfrac{\mathcal{E}}{N_0}\left(\dfrac{\gamma_0^2 c^2}{4}+D^2\overline{F^2}+\gamma_0 c D\overline{F'^2}\right)}. \tag{3.15}
$$

It covers the estimator that is locally unbiased for all the values of the nuisance parameter X. This kind of estimator may not be achieved, since for extremely low signal strengths, the receiver can not detect the signal any more [71]. The term

$1/[\mathbf{I}_D(\boldsymbol{\theta})]_{11}$ in MCB(D) is the bound for the estimator with perfect knowledge of X.

The fourth bound is the extended MCB (EMCB) [69]: EMCB(D) $= E_X\left\{\left[\mathbf{I}_D^{-1}(\boldsymbol{\theta})\right]_{11}\right\}$, which is

$$\mathrm{EMCB}(D) = E_X\left[\frac{1}{X^2}\right] \frac{c^2 D^{\gamma_0}}{2\widetilde{k_0}^2 \frac{\mathcal{E}}{N_0}\left\{\overline{F^2} - (\overline{F'^2})^2\right\}}. \tag{3.16}$$

It is the average over X of the joint estimation bound, which assumes X is an unknown deterministic parameter in each realization.

All the above bounds are related to each other as follows

$$\mathrm{CRB}(D) \geq \mathrm{HCRB}(D) \geq \mathrm{MCRB}(D), \tag{3.17}$$

$$\mathrm{EMCB}(D) \geq \mathrm{MCB}(D) \geq \mathrm{MCRB}(D). \tag{3.18}$$

The order (3.17) has already been proved in [72] [73]. The first inequality indicates that the CRB(D) (the true CRB for an AWGN channel with attenuation and shadowing) applied directly to D using a marginal probability density function is tighter than the HCRB(D). There is no performance improvement when estimating more parameters in the given system. However, due to calculation difficulties as we mentioned before, there is no closed form CRB(D) available. When we have $\overline{F'^2} = 0$, for UWB signals, in the relevant ranges, we would observe HCRB(D) $\approx$ MCRB(D), since $\gamma_0^2 c^2/4 \ll D^2\overline{F^2}$. For NB signals, we expect to see that the MCRB(D) is looser than the HCRB(D). The order (3.18) is also verified in [69]. The inequality $[\mathbf{I}_D^{-1}(\boldsymbol{\theta})]_{11} \geq 1/[\mathbf{I}_D(\boldsymbol{\theta})]_{11}$ always holds, which confirms the inequality EMCB(D) $\geq$ MCB(D). Again relying on $\overline{F'^2} = 0$, for UWB signals, we would have MCB(D) $\approx$ EMCB(D). For NB signals, differences will be obvious.

Table 3.2 collects the ranging accuracy obtained by different bounds for UWB signals and NB signals. The parameters are set the same values as in the last section. The shadowing parameter X follows a lognormal distribution $20\log_{10}X \sim \mathcal{N}(0, \sigma_x^2)$. The upper part of Table 3.2 is for UWB signals, while the lower part is for NB signals. The MCRB(D) is independent of the shadowing effects, due to the channel energy normalization. It is equal to the CRB in an AWGN channel only with attenuation. The order EMCB(D) $\approx$ MCB(D) $\geq$ HCRB(D) $\approx$ MCRB(D) is established from Table 3.2 for UWB signals in the relevant ranges similarly as we discussed in the last paragraph. It also indicates that the prior knowledge of X helps range estimation. As the shadowing effects increase, the estimators perform worse regardless of the prior knowledge of shadowing. For different σ_x, the HCRB(D) and the MCRB(D) remain the same for UWB signals. That's because we normalize the average channel energy, $\widetilde{k_0}E[X^2]$ is a constant. For NB signals, the differences between the HCRB(D) and the MCRB(D), as well

	$\sigma_{HCRB}(\mathrm{m})$	$\sigma_{MCRB}(\mathrm{m})$	$\sigma_{MCB}(\mathrm{m})$	$\sigma_{EMCB}(\mathrm{m})$
$D(\mathrm{m})$	Shadowing $\sigma_x = 3\mathrm{dB}$, UWB signals, $3.1\mathrm{GHz} \sim 10.6\mathrm{GHz}$			
1	$9.0788 \cdot 10^{-8}$	$9.0786 \cdot 10^{-8}$	$1.0089 \cdot 10^{-7}$	$1.0090 \cdot 10^{-7}$
3	$2.7236 \cdot 10^{-7}$	$2.7236 \cdot 10^{-7}$	$3.0269 \cdot 10^{-7}$	$3.0269 \cdot 10^{-7}$
10	$9.0788 \cdot 10^{-7}$	$9.0788 \cdot 10^{-7}$	$1.0090 \cdot 10^{-6}$	$1.0090 \cdot 10^{-6}$
$D(\mathrm{m})$	Shadowing $\sigma_x = 6\mathrm{dB}$, UWB signals, $3.1\mathrm{GHz} \sim 10.6\mathrm{GHz}$			
1	$9.0788 \cdot 10^{-8}$	$9.0786 \cdot 10^{-8}$	$1.3822 \cdot 10^{-7}$	$1.3822 \cdot 10^{-7}$
3	$2.7236 \cdot 10^{-7}$	$2.7236 \cdot 10^{-7}$	$4.1466 \cdot 10^{-7}$	$4.1466 \cdot 10^{-7}$
10	$9.0788 \cdot 10^{-7}$	$9.0788 \cdot 10^{-7}$	$1.3822 \cdot 10^{-6}$	$1.3822 \cdot 10^{-6}$
$D(\mathrm{m})$	Shadowing $\sigma_x = 3\mathrm{dB}$, NB signals, $0\mathrm{MHz} \sim 5\mathrm{MHz}$			
10	$2.5908 \cdot 10^{-2}$	$1.4317 \cdot 10^{-2}$	$1.5910 \cdot 10^{-2}$	$3.0750 \cdot 10^{-2}$
15	$4.0265 \cdot 10^{-2}$	$2.7884 \cdot 10^{-2}$	$3.0986 \cdot 10^{-2}$	$4.6125 \cdot 10^{-2}$
20	$5.4390 \cdot 10^{-2}$	$4.2648 \cdot 10^{-2}$	$4.7393 \cdot 10^{-2}$	$6.1500 \cdot 10^{-2}$
$D(\mathrm{m})$	Shadowing $\sigma_x = 6\mathrm{dB}$, NB signals, $0\mathrm{MHz} \sim 5\mathrm{MHz}$			
10	$2.2456 \cdot 10^{-2}$	$1.4317 \cdot 10^{-2}$	$2.1793 \cdot 10^{-2}$	$4.2121 \cdot 10^{-2}$
15	$3.7419 \cdot 10^{-2}$	$2.7884 \cdot 10^{-2}$	$4.2445 \cdot 10^{-2}$	$6.3182 \cdot 10^{-2}$
20	$5.2070 \cdot 10^{-2}$	$4.2648 \cdot 10^{-2}$	$6.4919 \cdot 10^{-2}$	$8.4242 \cdot 10^{-2}$

Table 3.2: Theoretical ranging accuracy for an AWGN channel with attenuation and shadowing $\sigma = \sqrt{\mathrm{CRB}}$

as between the MCB(D) and the EMCB(D) are obvious for the relevant ranges. The order EMCB(D) > HCRB(D) still holds. The HCRB(D) benefits from the prior knowledge of X. When $\sigma_x = 3\mathrm{dB}$, the HCRB(D) is larger than the MCB(D) around the critical distance. However, this relationship does not retain, when $\sigma_x = 6\mathrm{dB}$. As the shadowing effects become more serious, the MCB(D) becomes larger even if we know the exact value of X as assumed by the bound MCB(D). This is due to the fact that it bounds the average performance of a kind of estimator, which relies on the instantaneous received energy. Its performance is unfavorable, when the instantaneous received energy is low. On the other hand, the HCRB(D) is smaller when the shadowing is more severe and the MCRB(D) remains the same the under the condition that the average channel energy is normalized. The prior knowledge of X helps out when the instantaneous received energy is very low. In summary, it is important to take the prior knowledge into account to handle the shadowing effects.

Now let us investigate some estimators for X and D. There are two different ways to estimate X and D depending on whether to employ the prior knowledge of X in the estimation procedure. Method 1: If both X and D are assumed as unknown deterministic parameters, then the classic MLEs are derived. Method 2: if X is assumed as a random variable and $p(X)$ is known, D is assumed as an unknown deterministic parameter, then the Bayesian estimator is employed.

Method 1: we would like to find the X and the D that maximize $p(\mathbf{r}|X; D)$,

$$\frac{\partial \ln p(\mathbf{r}|X; D)}{\partial X} \;=\; 0, \qquad\qquad (3.19)$$

$$\frac{\partial \ln p(\mathbf{r}|X; D)}{\partial D} \;=\; 0. \qquad\qquad (3.20)$$

Solving (3.19), we obtain $\hat{X} = \dfrac{D^{\frac{\gamma_0}{2}}\mathcal{E}_{rs}(D)}{\widetilde{k_0}\mathcal{E}}$. Inserting $\hat{X}$ into (3.20) says that the estimation of D should satisfy the following equation

$$\int_0^{T_o} r(t)\frac{\partial s(t-\tau)}{\partial \tau}dt + \overline{F'^2}\mathcal{E}_{rs}(D) = 0. \tag{3.21}$$

Define $\overline{F'_{rs}{}^2}(D) = \int_0^{T_o} r(t)\frac{\partial s(t-\tau)}{\partial \tau}dt \Big/ \mathcal{E}_{rs}(D)$ conditioned on $\mathcal{E}_{rs}(D) \neq 0$. Then the MLE of D is equivalent to

$$\hat{D} = \operatorname*{argmin}_{0<D<cT_o} \left| \overline{F'_{rs}{}^2}(D) \right|. \tag{3.22}$$

In reality, we do not check the condition $\mathcal{E}_{rs}(D) \neq 0$, but test whether $\mathcal{E}_{rs}(D)$ is above the noise floor. If it is true, then we conclude that $\mathcal{E}_{rs}(D) \neq 0$ is satisfied. Otherwise, $\mathcal{E}_{rs}(D)$ equals zero and we have to check other D candidates. The average performance of this estimator can asymptotically approach $\text{EMCB}(D)$.

Method 2: In this case, we have prior knowledge of X, and a Bayesian estimator can be employed. The well-known minimum mean square error (MMSE) estimator is first employed $\hat{X} = E(X|\mathbf{r}; D) = \int Xp(X|\mathbf{r}; D)dX$, where $p(X|\mathbf{r}; D) = p(\mathbf{r}|X; D)p(X)\Big/ \int p(\mathbf{r}|X; D)p(X)dX$. The integration in the denominator prevents us from finding a closed form. Hence, we resort to a maximum a posteriori (MAP) estimator, which boils down to finding the X that maximizes $p(X|\mathbf{r}; D)$. It is equivalent to maximizing $p(\mathbf{r}|X; D)p(X)$ or $\ln(p(\mathbf{r}|X; D)p(X))$. Then, the joint estimation of X and D is

$$\begin{aligned}
[\hat{X}, \hat{D}] &= \operatorname*{argmax}_{X, 0<D<cT_o} \{\ln(p(\mathbf{r}|X; D)p(X))\} \\
&= \operatorname*{argmin}_{X, 0<D<cT_o} \{J(X, D)\}.
\end{aligned} \tag{3.23}$$

where $J(X, D) = g_1(X, D) + g_2(X)$,

$$\begin{aligned}
g_1(X, D) &= \frac{\mathcal{E}}{N_0}\left(\frac{\widetilde{k_0}X}{D^{\frac{\gamma_0}{2}}} - \frac{\mathcal{E}_{rs}(D)}{\mathcal{E}}\right)^2 - \frac{\mathcal{E}_{rs}^2(D)}{N_0\mathcal{E}}, \\
g_2(X) &= \frac{200}{\sigma_x^2\ln^2 10}\left(\ln X + \frac{\sigma_x^2\ln^2 10}{400}\right)^2 - \frac{\sigma_x^2\ln^2 10}{800}.
\end{aligned}$$

Method 2 is much more complicated than Method 1, since it has to execute a two-dimensional search compared to a one-dimensional search in Method 1. The performance limitation of Method 1 and Method 2 for range estimation can be indicated by the bound $\text{EMCB}(D)$ and the bound $\text{HCRB}(D)$, respectively.

3.4   Conclusions

In this chapter, we have investigated the accuracy for range estimation by a new method, which exploits the range information in both the amplitude and the time delay of the received signal. Several bounds are derived not only in an AWGN channel with attenuation, but also in an AWGN channel with attenuation and shadowing. For UWB signals, the new method does not show obvious benefits compared to an estimate exploring the range information in the delay only. However, it indeed improves the ranging accuracy using NB signals for ranges smaller than a threshold. Furthermore, taking the prior knowledge of the shadowing effects into account lowers the bounds for range estimation.

Remark that this chapter has focused on the relations among the bounds, and not their concrete values. In practice, it is always difficult to develop an optimal estimator to approach the bound. Many error sources besides noises, such as sampling effects, dense multipaths and unknown channel models, can dramatically increase gap between the theoretical bound and the achievable performance of practical estimators. In the next chapter, we will focus on a practical method for TOA-based range estimation using UWB signals in multipath environments.

3.A Derivation of CRB(D)(3.4)

Plugging (3.2) into (3.3), we derive the CRB for D as follows:

$$
\begin{aligned}
\mathrm{CRB}(D) &= -\left\{ E_{\mathbf{r}}\left[\frac{\partial^2}{\partial D^2}\ln p(\mathbf{r};D)\right]\right\}^{-1} \\
&= \sigma^2 \left\{ \sum_{n=0}^{N-1}\left(\frac{\partial k_0 D^{-\frac{\gamma_0}{2}}\mathbf{s}_D}{\partial D}\right)^2 \right\}^{-1} \\
&= \sigma^2 \left\{ \frac{1}{T_s}\int_0^{T_o}\left(\frac{\partial}{\partial D}k_0 D^{-\frac{\gamma_0}{2}}s(t-\tau)\right)^2 dt \right\}^{-1} \\
&= \frac{N_0}{2}\left\{ \int_0^{T_o}\left(\frac{\partial}{\partial D}k_0 D^{-\frac{\gamma_0}{2}}s(t-\tau)\right)^2 dt \right\}^{-1} \\
&= \frac{N_0}{2}\left\{ \int_0^{T_o}\left(-\frac{\gamma_0}{2}k_0 D^{-\frac{\gamma_0}{2}-1}s(t-\tau) + c^{-1}k_0 D^{-\frac{\gamma_0}{2}}\frac{\partial s(t-\tau)}{\partial \tau}\right)^2 dt \right\}^{-1} \\
&= \frac{N_0}{2}\left\{ \int_0^{T_o}\left(\frac{\gamma_0^2}{4}k_0^2 D^{-\gamma_0-2}s^2(t-\tau) + c^{-2}k_0^2 D^{-\gamma_0}\left(\frac{\partial s(t-\tau)}{\partial \tau}\right)^2 \right. \right. \\
&\qquad \left. \left. -\gamma_0 c^{-1}k_0^2 D^{-\gamma_0-1}s(t-\tau)\frac{\partial s(t-\tau)}{\partial \tau}\right) dt \right\}^{-1} \\
&= \frac{N_0}{2}\left\{ \int_0^{T_o}\left(\frac{\gamma_0^2}{4}k_0^2 D^{-\gamma_0-2}s^2(t-\tau) + c^{-2}k_0^2 D^{-\gamma_0}\left(-\frac{\partial s(t-\tau)}{\partial (t-\tau)}\right)^2 \right. \right. \\
&\qquad \left. \left. +\gamma_0 c^{-1}k_0^2 D^{-\gamma_0-1}s(t-\tau)\frac{\partial s(t-\tau)}{\partial (t-\tau)}\right) dt \right\}^{-1} \\
&= \frac{N_0}{2}\frac{c^2 D^{\gamma_0+2}}{k_0^2}\left\{ \int_{-\tau}^{T_o-\tau}\left(\frac{\gamma_0^2 c^2}{4}s^2(t) + D^2\left(\frac{\partial s(t)}{\partial t}\right)^2 \right. \right. \\
&\qquad \left. \left. +\gamma_0 c D s(t)\frac{\partial s(t)}{\partial t}\right) dt \right\}^{-1}
\end{aligned}
\tag{3.24}
$$

Using the definitions of $\mathcal{E}$, $\overline{F^2}$ and $\overline{F'^2}$, we can rewrite (3.24) as (3.4).

3.B   Derivation of $\mathbf{I}_D(\,\theta)$

We derive each element of $\mathbf{I}_D(\boldsymbol{\theta})$:

$$
\begin{aligned}
[\mathbf{I}_D(\boldsymbol{\theta})]_{11} &= -E_{\mathbf{r}|\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial D^2}\ln p(\mathbf{r}|\boldsymbol{\theta})\right] \\
&= \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{\partial \widetilde{k}_0 D^{-\frac{\gamma_0}{2}}X\mathbf{s}_D}{\partial D}\right)^2 \\
&= \frac{2X^2}{N_0}\int_0^{T_o}\left(\frac{\partial}{\partial D}\widetilde{k}_0 D^{-\frac{\gamma_0}{2}}s(t-\tau)\right)^2 dt \\
&= \frac{2\widetilde{k}_0^2}{D^{\gamma_0}}\frac{X^2}{N_0}\int_{-\tau}^{T_o-\tau}\left(\frac{\gamma_0^2}{4D^2}s^2(t)+\frac{1}{c^2}\left(\frac{\partial s(t)}{\partial t}\right)^2+\frac{\gamma_0}{cD}s(t)\frac{\partial s(t)}{\partial t}\right) dt
\end{aligned}
$$

$$(3.25)$$

$$
\begin{aligned}
[\mathbf{I}_D(\boldsymbol{\theta})]_{12} &= -E_{\mathbf{r}|\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial D\partial X}\ln p(\mathbf{r}|\boldsymbol{\theta})\right] \\
&= \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{\partial \widetilde{k}_0 D^{-\frac{\gamma_0}{2}}X\mathbf{s}_D}{\partial D}\right)\left(\frac{\partial \widetilde{k}_0 D^{-\frac{\gamma_0}{2}}X\mathbf{s}_D}{\partial X}\right) \\
&= \frac{2}{N_0}X\int_0^{T_o}\widetilde{k}_0 D^{-\frac{\gamma_0}{2}}s(t-\tau) \\
&\qquad\left(-\frac{\gamma_0}{2}\widetilde{k}_0 D^{-\frac{\gamma_0}{2}-1}s(t-\tau)-c^{-1}\widetilde{k}_0 D^{-\frac{\gamma_0}{2}}\frac{\partial s(t-\tau)}{\partial(t-\tau)}\right) dt \\
&= -\frac{2\widetilde{k}_0^2}{D^{\gamma_0}}\frac{X}{N_0}\int_{-\tau}^{T_o-\tau}\left(\frac{\gamma_0}{2D}s^2(t)+\frac{1}{c}s(t)\frac{\partial s(t)}{\partial t}\right) dt
\end{aligned}
$$

$$(3.26)$$

$$
\begin{aligned}
[\mathbf{I}_D(\boldsymbol{\theta})]_{22} &= -E_{\mathbf{r}|\boldsymbol{\theta}}\left[\frac{\partial^2}{\partial X^2}\ln p(\mathbf{r}|\boldsymbol{\theta})\right] \\
&= \frac{1}{\sigma^2}\sum_{n=0}^{N-1}\left(\frac{\partial \widetilde{k}_0 D^{-\frac{\gamma_0}{2}}X\mathbf{s}_D}{\partial X}\right)^2 \\
&= \frac{2\widetilde{k}_0^2}{D^{\gamma_0}}\frac{1}{N_0}\int_0^{T_o}s^2(t-\tau)dt
\end{aligned}
$$

$$(3.27)$$

Based on the above results and recalling $\mathcal{E}$, $\overline{F^2}$ and $\overline{F'^2}$, we can write $\mathbf{I}_D(\boldsymbol{\theta})$ as (3.13).

Time-of-Arrival Estimation by UWB Radios

4.1 Introduction

Ultra-wideband (UWB) impulse radio (IR) is a promising technology for high resolution time-of-arrival (TOA) estimation [1, 74–78]. It enables precise ranging and facilitates accurate positioning, which has a wide range of applications such as environment monitoring and control, target tracking, industrial quality control and emergency services [79–81]. From an estimate of the first arriving component of UWB IR in a dense multipath environment, we can estimate the TOA with high accuracy. In a line-of-sight (LOS) environment, the first path component is usually the strongest one and we can perform peak selection (PS) [77], whereas in a non-line-of-sight (NLOS) scenario, the first path component normally is not the strongest one, in which case we have to resort to other alternatives, such as jump back and search forward (JBSF) [77, 82].

Due to the large bandwidth of UWB IR, its multipath components are resolvable, which is benefit for accurate TOA estimation. However, for the same reason, the IR-UWB system requires a Nyquist sampling rate of several GHz, which is prohibitively high for practical implementation [15, 16]. Most ranging systems are based on Nyquist sampling in order to take full advantage of the large bandwidth to achieve high accuracy, such as the generalized maximum likelihood (ML) TOA estimator proposed in [74], and various TOA estimation strategies applied to the Nyquist-rate sampled output of a matched filter (MF) in [83]. Recently, more and more research has been devoted to sub-Nyquist TOA estimation due to its simplicity. The energy detector (ED) [41, 84, 85], the transmitted-reference (TR)

The results in this chapter appeared in [32]

receiver [41, 86] and the stored reference correlator [41] use the received signal itself, its delayed version and a noiseless template respectively as templates to generate energy samples at a sub-Nyquist sampling rate. Different kinds of TOA estimation techniques can be used to extract timing information from these energy samples, such as maximum energy sum selection (MESS), weighted MESS [41, 87], P-Max selection [83], threshold crossing (TC) [83, 84, 86, 88], consecutive threshold crossing [85], serial backward search (SBS) and serial backward search for multiple clusters (SBSMC) [82]. However, in all these schemes, the accuracy of TOA estimation is sacrificed for sub-Nyquist sampling. According to these previous research work, the high accuracy and the low sampling rate are regarded as conflict aspects. It seems impossible to achieve them in UWB systems at the same time.

In this chapter, we take the challenge to obtain a high resolution TOA estimate using UWB IR with low sampling rate. We make use of stroboscopic sampling, which is widely used in channel measurements [32, 89]. It can obtain an effective sampling rate as high as several GHz using a low-rate analog-to-digital converter (ADC) running at several tens or hundreds of MHz with the penalty of repetitively transmitting the same waveform. Since we have to transmit several identical pulses in order to collect the same samples as when transmitting one pulse sampled by an equivalent high sampling rate, the preamble for ranging is long. The longer preamble is the cost to obtain the high accuracy and the low sampling rate at the same time, which is still affordable for a UWB system in purpose of accurate ranging. Furthermore, whenever we consider TOA estimation for ranging, the clock always plays an essential role. Due to the randomness of the clocks in reality, the clock drift, which refers to the phenomenon where the clock does not run at the nominal frequency, becomes one of the main error sources [75, 90]. It does not only cause problems in ranging computation, but also in TOA estimation. The problems are more severe in a stroboscopic sampling system, since it needs more time to collect sufficient samples. Although a symmetric double-sided two-way ranging (SDS-TWR) protocol as presented in [75, 76] can reduce the ranging error due to clock drift significantly by relating the drift to the difference of the processing times at the two devices instead of the processing time at one device, it cannot calibrate for the TOA estimation error caused by the clock drift. The protocol works under the assumption that the TOA is estimated correctly. A delay-locked-loop (DLL) circuit is used in [90] to lock the clock frequency of the ranging responder to the estimated clock of the ranging initiator. However, a DLL may not be appropriate for a stroboscopic sampling scheme, since the responder and the initiator employ different clock frequencies. A trellis-based ML crystal drift estimator is introduced in [91, 92] to solve this problem. However, the system sampling rate is still as high as several GHz. Furthermore, it does not consider the code mismatch due to the drift and the unknown TOA. Therefore, we design a ranging preamble, solve the code mismatch problem and apply an ML estimator (MLE) to determine the clock drift in our stroboscopic sampling IR-UWB system. After the drift calibration, various TOA estimation methods are investigated to choose proper ones considering performance and cost. Consequently, we

calibrate for the timing error caused by the clock drift and achieve an accurate TOA estimate with low sampling rate. Moreover, the Cramèr-Rao bound (CRB) is derived to indicate a theoretical performance limit for joint estimation of the clock drift ratio and the parameters of the multipath components, which is different from the conventional CRBs only taking the multipath components into account [1, 77, 93, 94].

The rest of this chapter is organized as follows. In Section 4.2, we will first introduce the stroboscopic sampling principle. The clock drift is taken into account in the system model and the preamble is designed to facilitate the drift calibration. In Section 4.3, we propose a method to accurately estimate the TOA with drift calibration. In Section 4.4, we derive the Cramèr-Rao bound (CRB) for joint estimation of the clock drift ratio and the parameters of the multipath components. Simulation results are shown in Section 4.5. Conclusions are drawn at the end of the chapter.

4.2 System model

The preamble for ranging is composed of many frames. Each frame period T_f holds one pulse. We design T_f to be larger than the delay spread of the channel in order to avoid inter-frame interference (IFI) and reserves some margin to accommodate noise only to facilitate the TOA estimation later. The receiver employs a front-end filter to select the band of interest as shown in Fig. 4.1. The impulse response of the front-end filter does not have to be the received pulse shape, which is unknown due to the distortions caused by the channel and the antennas. For instance, it can be the transmitted pulse. In general, we specify the filter in the frequency domain in order to capture most of the signal energy in the band of interest. The bandwidth B of the front-end filter is quite large because of the bandwidth of the UWB signal. Hence, the Nyquist sampling rate becomes prohibitively high. Therefore, we resort to stroboscopic sampling [89] to sample the output of the front-end filter at rate $1/T_{\mathrm{sam}}$, which is much smaller than its Nyquist sampling rate $2B$, i.e., $T_{\mathrm{sam}} \geq 1/(2B)$. Consequently, each frame produces $\lfloor T_f/T_{\mathrm{sam}} \rfloor$ or $\lceil T_f/T_{\mathrm{sam}} \rceil$ samples. Since we apply stroboscopic sampling, several identical frames have to be transmitted in order to collect a sufficient number of samples that are equivalent to those obtained by sampling one frame at a higher rate. We define the equivalent high sampling rate as $1/T_b$, which satisfies the condition $T_b \leq 1/(2B)$ in order to prevent frequency aliasing. The resolution of TOA estimation, which is the smallest resolvable time difference, is determined by the equivalent sampling period T_b ($\leq T_{\mathrm{sam}}$). The sample vector $\mathbf{x}$ in Fig. 4.1 is the input of a digital signal processing (DSP) block. The DSP block would accomplish TOA estimation with clock drift calibration. We remark that the analog-to-digital convertor (ADC) used in stroboscopic sampling is comparable to one subconvertor of a time-interleaved ADC [95], which employs multiple subconvertors running at low rate in parallel to achieve an overall high
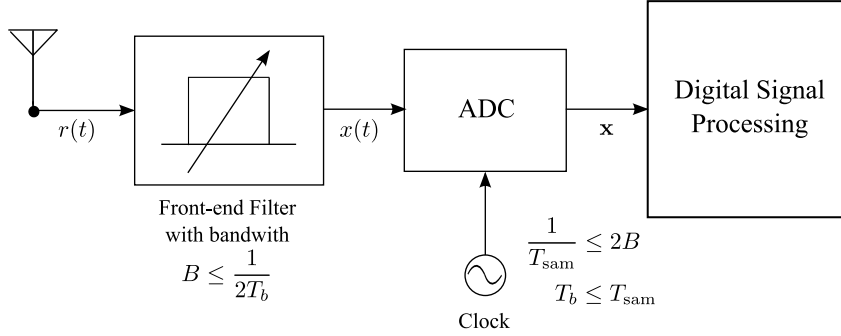
Figure 4.1: The receiver's analog front-end. The output of the front-end filter is sampled at rate $1/T_{\mathrm{sam}}$, which is smaller than its Nyquist sampling rate $2B$.

sampling rate. Ideally, the resolution of TOA estimation using time-interleaved ADC and using stroboscopic sampling should be the same. They make different trade-offs between the time and the hardware cost. In general, the ADC starts with a sample-and-hold (S/H) operation followed by digital quantization. In the case of stroboscopic sampling, the S/H circuit has to be fast enough to follow the change of the UWB signal [96]. The aperture jitter or aperture uncertainty of the ADC [97], which refers to the random variation of the sampling instant, should be negligible compared to the equivalent sampling period T_b [98]. For example, if the targeted equivalent sampling period T_b is 1 ns, the aperture jitter should be limited to a few picoseconds, which is feasible as reported in [15]. Moreover, since the stroboscopic sampling only employs one ADC, it would not suffer from the subconvertor mismatch problem, which imposes a big challenge to the time-interleaved ADC [99].

The relationships among T_{sam}, T_f, and T_b are given as follows

$$
\begin{aligned}
T_{\mathrm{sam}} &= mT_b, & (4.1)\\
T_f &= (mP + q)T_b, & (4.2)\\
\frac{T_f}{T_b} &= m\frac{T_f}{T_{sam}}, & (4.3)
\end{aligned}
$$

where $m \geq 1$ is the sampling-rate gain, $P = \lfloor T_f/T_{\mathrm{sam}} \rfloor$ is the minimum number of samples collected from one frame through stroboscopic sampling and $m > q > 0$. These parameters are all integers. Moreover, m and $mP + q$ should be relatively prime. Under the condition that $m = 1$ and $q = 0$, the system becomes a Nyquist sampling system. When designing the sampling-rate gain m, we would like it to be as large as possible to lower the sampling rate. On the other hand, it has to be small to shorten the preamble to save transmission energy. So there clearly is a design trade-off. Using m frames to collect $mP + q$ samples is equivalent to
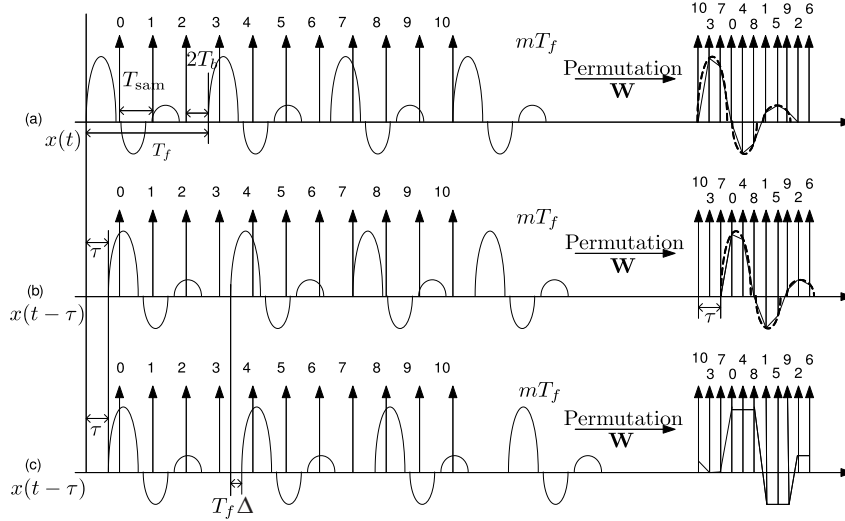
Figure 4.2: The noiseless output of the front-end filter at the receiver at rate $1/T_{\mathrm{sam}}$. (a) The ideal case, no drift and $\tau = 0$, (b) no drift and $\tau \neq 0$, and (c) with drift and $\tau \neq 0$.

sampling one frame at rate $1/T_b$. Fig. 4.2(a) shows an example. The waveform $x(t)$ represents the output of the front-end filter. In the example, $m = 3$, $P = 3$ and $q = 2$, which leads to $T_f = (mP+q)T_b = 11T_b$. We collect 11 samples in a time duration of $3T_f$. These samples are equivalent to those obtained by sampling one frame at the output of the front-end filter at rate $1/T_b$, and permuting as shown at the right side of Fig. 4.2(a).

The relative clock drift between the transmitter and the receiver violates the relation in (4.2). That is because T_f is with respect to the transmitter, while T_{sam} is with respect to the receiver. When we apply stroboscopic sampling at the receiver, we require a long preamble in order to obtain a sufficient number of samples. But since the relative clock drift ratio can be as large as 80 ppm [75], the drift of the preamble can lead to serious problems in TOA estimation. Therefore, we have to calibrate for the clock drift at the receiver for accurate ranging. Figs. 4.2(b) and 4.2(c) indicate examples without and with clock drift, respectively, where Δ is the relative clock drift ratio and τ is the TOA. In Fig. 4.2(b), the timing information can be retrieved after permuting the sample sequence, whereas in Fig. 4.2(c) the original waveform can not be regained due to the clock drift, and the timing information is lost.

Assuming the clock drift ratio remains constant, the clock drift linearly increases with time. In order to suppress the noise, we define a group of frames as a cluster according to the prior knowledge of the maximum drift ratio. The duration of a cluster is assumed smaller than the minimum time period required

to observe a drift of T_b. Therefore, relation (4.2) is roughly maintained within the cluster. Recalling that $1/T_b$ is larger than $2B$, a maximum drift of T_b within a cluster is much smaller than the width of the pulse, which is one of the parameters to decide the bandwidth of the UWB system. Hence, frame samples can be averaged within the cluster without severe pulse mismatch. The cluster period is defined as $N_f T_f$, where $N_f = mM$ and $M > 0$ is the processing gain. It has to satisfy the following condition [91, 92]

$$N_f T_f \Delta_{\max} = mMT_f \Delta_{\max} \leq T_b, \tag{4.4}$$

where $\Delta_{\max}$ is the maximum clock drift ratio at the transmitter relative to the receiver. For instance, if the targeted T_b is 1 ns, $T_f = 150$ ns, and $\Delta_{\max}$ is 80 ppm, then $N_f \leq 83$, which means $mM \leq 83$. A proper choice of m could be 7, which indicates a sampling rate of $1/T_{\mathrm{sam}} = 142.9$ MHz or a sampling period of $T_{\mathrm{sam}} = 7$ ns. As a result, the largest processing gain M can be 12. The outcome of the cluster averaging are $mP + q$ samples of one frame.

In order to achieve TOA estimation, we could design the preamble to be composed of several segments, each of which is dedicated to serve a different purpose, such as signal detection, coarse synchronization and fine synchronization. Each segment could have a different structure to facilitate its task. The structural design of the whole preamble is out of the scope of this chapter. We assume coarse synchronization has already been carried out. More specifically, we assume that the TOA τ is in the range of one frame period with respect to the receiver, i.e., $\tau \in [0, (1 + \Delta)T_f)$, where $(1 + \Delta)T_f \approx T_f$, since T_f is only several hundreds of ns and $\Delta \leq 80$ ppm. Therefore, we only concentrate on designing the preamble for the fine synchronization stage where τ is estimated. We assign a code to each cluster instead of each frame in order to avoid code mismatch during the averaging due to the unknown τ and the clock drift, which is not considered in [91, 92]. We remark that the purpose of the code here is for spectrum smoothing and multiuser accessing, not for drift calibration, which can still work without any code assignment. Therefore, we have not considered the code design in this chapter. Based on the analysis above, the structure of the transmitted preamble is shown in Fig. 4.3(a). The preamble is composed of N_c clusters, where every cluster is made up of N_f frames, each one containing one pulse. The transmitted signal can thus be written as

$$s(t) = \sum_{i=0}^{N_c-1} c_i b(t - iN_f T_f), \tag{4.5}$$

where $b(t) = \sum_{j=0}^{N_f-1} p(t - jT_f)$, c_i is the cluster code and $p(t)$ is the transmitted pulse shape of a very narrow width. Figs. 4.3(b) and 4.3(c) show the noiseless received preamble through an additive white Gaussian noise (AWGN) channel with unknown τ and different clock drifts. As we can observe from the figures,
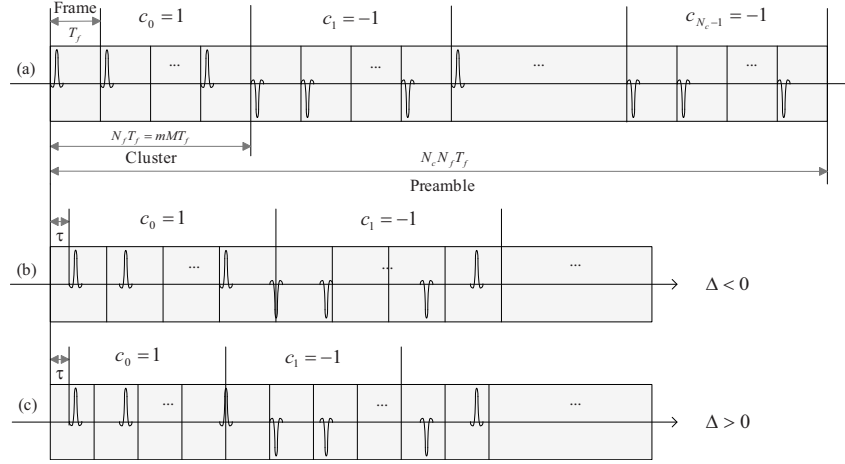
Figure 4.3: The structure of (a) the preamble, (b) the received preamble with negative drift, and (c) the received preamble with positive drift.

there is a code mismatch due to the unknown τ and the clock drift. The last several frames of the clusters in Fig. 4.3(b) are mismatched due to the unknown τ and the negative drift, while the first several frames of the clusters in Fig. 4.3(c) are mismatched as a result of the unknown τ and the positive drift. We cut off the first and last m frames of each cluster to get rid of the code mismatch in the averaging process at the price of reducing the processing gain from M to $M - 2$. As illustrated in the previous example, an 80 ppm drift ratio may cause an 80 ns clock drift for a 1 ms preamble. Due to this drift and the unknown $\tau \in [0, 150 \text{ ns})$, the timing offset at the last frame of the preamble will be in the range of $[\tau - 80 \text{ ns}, \tau + 80 \text{ ns})$, which leads to a timing offset range of $[-80 \text{ ns}, 230 \text{ ns}]$. The m frames we omit provide a guard time, which is larger than the timing offset range. Hence, it can prevent code mismatch. When $m = 1$, which indicates Nyquist sampling, however, the timing offset range is larger than the guard time, and the first/last frame still suffers from a code mismatch even after frame removal. Nevertheless, since we average lots of frames over the cluster in case of Nyquist sampling ($m = 1$), a code mismatch in one frame does not introduce a big influence. The received preamble is

$$r(t) \quad = \quad \sum_{l=0}^{L-1} \alpha_l s(t - \tau_{l,0} - \tau, \Delta) + n(t), \tag{4.6}$$

where $n(t)$ is the zero-mean AWGN with double sided power spectral density

$N_0/2$,

$$s(t, \Delta) = \sum_{i=0}^{N_c-1} c_i b(t - iN_f T_f(1+\Delta), \Delta), \qquad (4.7)$$

$$b(t, \Delta) = \sum_{j=0}^{N_f-1} p(t - jT_f(1+\Delta)), \qquad (4.8)$$

Δ is the clock drift ratio, L indicates the number of multipath components, and α_l and $\tau_{l,0}$ represent the amplitude and the relative time delay of the lth path with respect to the first path, respectively. Note that $\tau_{l,0} = \tau_l - \tau_0$, where τ_l is the multipath delay and $\tau_0 = \tau$. We remark here the frame period at the transmitter is T_f, and the one at the receiver is $T_f(1+\Delta)$. With the received signal $r(t)$, the output of the front-end filter sampled using stroboscopic sampling is

$$x(nT_{\text{sam}}) = \int_{-\infty}^{+\infty} r(t)g(nT_{\text{sam}} - t)dt, \quad n = 0, 1, \ldots, MN_cL_f, \qquad (4.9)$$

where $g(t)$ is the impulse response of the front-end filter, whose bandwidth is large enough to include the band of interest, and $L_f = T_f/T_b = mP + q$ is the frame length in terms of the number of samples at rate $1/T_b$. Let us introduce $\mathbf{x}_k = \left[x((kL_f)T_{\text{sam}}) \ x((kL_f + 1)T_{\text{sam}}) \ \ldots \ x((kL_f + L_f - 1)T_{\text{sam}}) \right]^T$, $k = 0, 1, \ldots, N_cM - 1$, which is an L_f-long sample vector for the kth equivalent frame. Notice that $\mathbf{x}_k$ is the result of sampling m frames at rate $1/T_{\text{sam}}$ at the receiver. Every M sample vectors are grouped as a cluster. We exclude the first and last sample vectors in each cluster to get rid of the code mismatch in the averaging process. The results are collected in a data matrix $\mathbf{X}$ of size $L_f \times N_c$, which is given by

$$\mathbf{X} = \frac{1}{M-2} \left[\sum_{i=1}^{M-2} \mathbf{x}_i \ \sum_{i=1}^{M-2} \mathbf{x}_{M+i} \ \ldots \ \sum_{i=1}^{M-2} \mathbf{x}_{(N_c-1)M+i} \right], \qquad (4.10)$$

where each column of $\mathbf{X}$ contains its own specific code.

4.3   TOA estimation with clock drift calibration

4.3.1   Recovery from stroboscopic sampling

Due to the stroboscopic effects, we have to permute all the averaged frame samples in each column of $\mathbf{X}$ before we calibrate for the drift and estimate the TOA τ. The adjacent T_{sam}-spaced samples obtained by stroboscopic sampling are not the adjacent T_b-spaced samples in the equivalent high sampling rate scheme as shown in Fig. 4.2(a). The maximum drift observed in a cluster is T_b, which is much

smaller than the sample spacing T_{sam} in the stroboscopic sampling scheme, and exactly equal to the sample spacing in the equivalent high sampling rate scheme. Therefore, we have to recover the equivalent high sampling rate sequence before drift calibration to appropriately represent the drift between the adjacent clusters. Recall that m is the sampling-rate gain. According to (4.2), we define a permutation matrix $\mathbf{W}$ of size $L_f \times L_f$ with first column $[\mathbf{W}]_{:,1} = [\mathbf{0}_{m-1}^T \ 1 \ \mathbf{0}_{L_f-m}^T]^T$ and every other column a circulant shift of the previous column. This means that

$$[\mathbf{W}]_{:,i+1} = \mathrm{circshift}([\mathbf{W}]_{:,i}, m), i = 1, \ldots, L_f - 1, \qquad (4.11)$$

where $\mathrm{circshift}(\mathbf{a}, n)$ circularly shifts the values in the vector $\mathbf{a}$ by $|n|$ elements (down if $n > 0$ and up if $n < 0$). The rearrangement is accomplished by

$$\mathbf{X}_o = \mathbf{W}\mathbf{X}, \qquad (4.12)$$

where each column of $\mathbf{X}_o$ collects the permuted averages for each cluster. Now, the equivalent sample spacing in $\mathbf{X}_o$ is T_b.

4.3.2 An estimator of clock drift

In order to estimate the TOA, we would like to use all the data samples in $\mathbf{X}_o$. This allows us to obtain an averaged sample vector over all the clusters and thereby reduces the noise. However, due to the clock drift, the equivalent frame waveforms do not align with each other. We have to calibrate for the drift before TOA estimation. Let us define the row index of the data matrix $\mathbf{X}_o$ as the frame phase, similar to the pulse repetition period (PRP) phase in [91, 92] (see an example in Fig. 4.4). We recall that the maximum drift accumulated over a cluster duration is T_b, and the difference between two adjacent phase is also at most T_b. We have assumed that the accumulated drift, which is the total drift observed over the duration from the beginning of the preamble to the interested time point, is zero at the start of the preamble. The frame phase of a cluster may correspond to the same or an adjacent phase in the next cluster due to the accumulated drift. This kind of correspondence is called the transition between frame phases. The drift estimation traces the correct transition path of the frame phase within the duration of the preamble. A transition takes place between two contiguous clusters, when the accumulated drift increases by T_b. The transition path pattern is generated by quantizing the accumulated drift over clusters. The quantization step size is T_b, and a ceiling quantization is employed. Every frame phase has the same set of transition paths. We remark that the exact value of the drift ratio is not the main concern, but the transition path is the target of the drift estimation. Based on this path, we can calibrate for the drift, obtain an averaged sample vector for the whole preamble, and then estimate the TOA. The resolution of the drift estimation is $T_b/((N_c - 1)N_f T_f)$. The total number of transition paths for each frame phase is $2N_c - 1$, where $N_c - 1$ is the path number

for positive drift (or negative drift), and one path is for the case without any drift. Therefore, the longer the preamble, the more accurate the drift estimation. Nevertheless, the complexity of the estimation would also increase as the number of transition paths increases.

Fig. 4.4 shows some examples of transition paths. In the example, $N_c = 5$ and $L_f = 13$. The dots represent the elements of the matrix $\mathbf{X}_o$. The spacing between contiguous samples is T_b. Based on Fig. 4.4, we reconfirm that the data matrix $\mathbf{X}$ can not be used directly. Some of the transition paths for frame phase 7 are shown. As mentioned before, we have assumed that the accumulate drift at the beginning of cluster 1 is zero. Path 1 indicates that we can observe a phase transition for every cluster, and the kth phase of the ith cluster transfers to the $(k-1)$th phase of the $(i+1)$th cluster. It reaches the negative maximum drift. In path 5, no clock drift is observed. Path 6 shows T_b drift over five clusters, which corresponds to a resolution of the drift ratio estimation given by $\Delta_{\max}/4$. We remark that path 6 is the only transition path for which we observe a positive drift of T_b over five clusters, according to the quantization rules and the assumption of the zero accumulated drift at the start of the preamble. The number of all possible transition paths for each frame phase is $2N_c - 1 = 9$. There are special cases we have to be careful with. For example, in path a for phase 2, there is a negative shift of one phase every cluster. Therefore, phase 1 of cluster 2 transfers to phase 13 of cluster 3 as shown by the dashed line with solid arrow in the figure. Meanwhile, path b for phase 10 describes a positive shift of one phase every cluster propagates in a similar way. The transition takes place circularly. Since we have excluded the code mismatch during the cluster averaging process by cutting off the first and last sample vectors, each column of $\mathbf{X_o}$ corresponds to its own specific code. However, the phase mismatch due to the clock drift still causes serious problems for TOA estimation.

All the transition paths for any phase are modeled in a matrix $\mathbf{\Lambda}$ of size $(2N_c - 1) \times N_c$. For instance, $\mathbf{\Lambda}$ for $N_c = 5$ is shown at the right side of Fig. 4.4. The path number in Fig. 4.4 corresponds to the row index of $\mathbf{\Lambda}$. Making use of the transition matrix and recalling the cluster code, we can estimate the transition path and calibrate for the clock drift. Since all the phases have the same set of transition paths, in order to be more robust to noise we perform an exhaustive search of the transition path that leads to an averaged frame waveform with the maximum energy, which can be regarded as the maximum likelihood (ML) estimator for the clock drift. We remark that instead of jointly considering drift calibration with other TOA estimation techniques, we decouple these two problems and solve them sequentially in order to obtain lower computation complexity. A simple joint approach is to estimate the clock drift and the strongest multipath component at the same time [32]. However, it requires a two-dimensional search. Moreover, in a NLOS scenario, the first multipath component is in general not the strongest one, so looking for the strongest path is not always the best option for TOA estimation. We define $j_{\max}$ as the row index of the selected path in $\mathbf{\Lambda}$. The problem we then
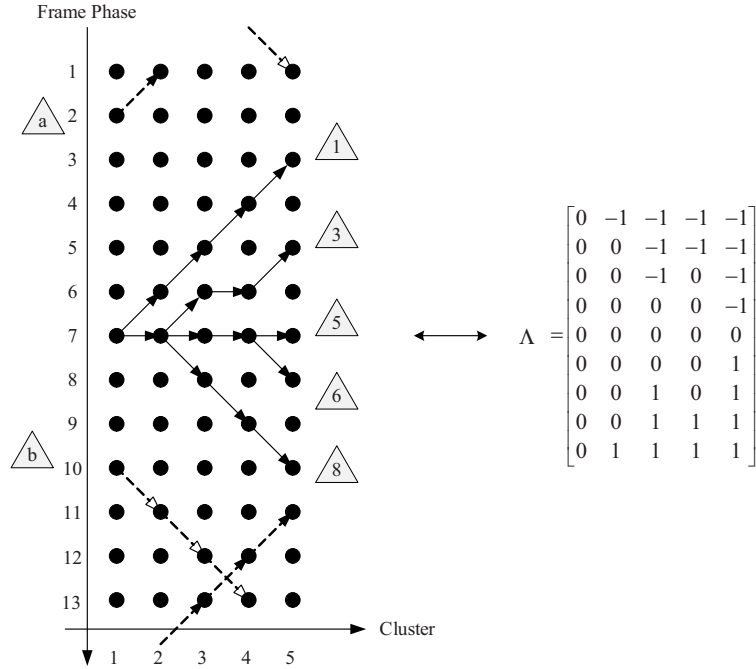
Figure 4.4: The diagram of the transition paths and the corresponding matrix $\mathbf{\Lambda}$ for $N_c = 5$ and $L_f = 13$.

would like to solve is

$$
j_{\max} = \arg \max_{j \in \{1, \ldots, 2N_c - 1\}} \sum_{k=1}^{L_f} \left| \sum_{n=1}^{N_c} c_{n-1} [\mathbf{X}_o]_{p(k,n,j),n} \right|^2, \qquad (4.13)
$$

$$
p(k,n,j) = \mathrm{mod}\left((k - 1 + \sum_{i=1}^{n} \mathbf{\Lambda}(j,i)), L_f\right) + 1, \qquad (4.14)
$$

where the function $\mathrm{mod}(a,b)$ computes a modulo b. It is equivalent to searching for the transition path, which results an averaged frame waveform with the maximum energy. The clock drift ratio corresponding to the selected transition path can be estimated as

$$
\hat{\Delta} = \frac{(j_{\max} - N_c)T_b}{(N_c - 1)N_f T_f}, \qquad (4.15)
$$

which is the ratio between the accumulated clock drift and the total observation duration. We remark that the complexity of the drift estimator depends on the

total searching space. In the scope of the thesis, we have not considered the complexity to decide the maximum value and the optimal searching strategy. For the clock drift estimator, in the worst case scenario, we carry out an exhaustive search over all the possible paths. According to (4.13) and (4.14), the number of operations required by the clock drift estimator is in the order of $O(L_f N_c(2N_c - 1))$.

4.3.3   TOA estimation

Until now, we have estimated the transition path for the clock drift using (4.13). Before we continue to discuss the TOA estimation strategies, we first average the related phases of $\mathbf{X}_o$ over the whole preamble according to the transition path $\mathbf{\Lambda}(j_{\max}, :)$ to calibrate for the drift and further mitigate the noise, and collect the outcomes in a sample vector $\mathbf{y}$, which is actually a circularly shifted version of the channel estimate. Its elements are computed as

$$[\mathbf{y}]_k = \frac{1}{N_c} \sum_{n=1}^{N_c} c_n [\mathbf{X}_o]_{p(k,n,j_{\max}),n}, k = 1, 2, \ldots, L_f. \tag{4.16}$$

The vector $\mathbf{y}$ is used for TOA estimation. As a result, the total processing gain for TOA estimation is $N_c(M - 2)$.

We define $k_{\max}$ as the index of the strongest multipath component, which is found by

$$k_{max} = \arg \max_{k \in \{1,2,\ldots,L_f\}} |[\mathbf{y}]_k|. \tag{4.17}$$

Thus, the TOA estimation according to the PS method [41,83] is given by $\hat{\tau}_p = T_b k_{\max} - \frac{T_b}{2}$. In a NLOS scenario, the problem to detect the leading edge arises, since the first path may not be the strongest one. Therefore, we have to resort to other alternatives. The challenge in TOA estimation is imposed by the unknown statistical properties of the channel.

There are various kinds of TOA estimation strategies as we mentioned in the introduction [77, 83, 88]. They are summarized in Table 4.1. An example of how these TOA strategies can be applied to $|\mathbf{y}|$ is shown in Fig. 4.5. The TOA estimation based on P-Max selection chooses the index of the first sample among the P largest samples $|[\mathbf{y}]_k|$. Its performance is sensitive to P and it may lock to the strongest cluster as shown in Fig. 4.5, in which we choose $P = 5$. The threshold crossing (TC) method [83, 84, 86, 88] selects the index of the first sample, whose absolute value exceeds the threshold. It is not appropriate to be applied in our case: as $\mathbf{y}$ is a circularly shifted version of the channel estimate, the TC method may regard the channel tail as the leading edge. An incorrect choice by the TC method is indicated in Fig. 4.5. The TOA estimation based on the serial backward search (SBS) method [77] first finds out the index of the
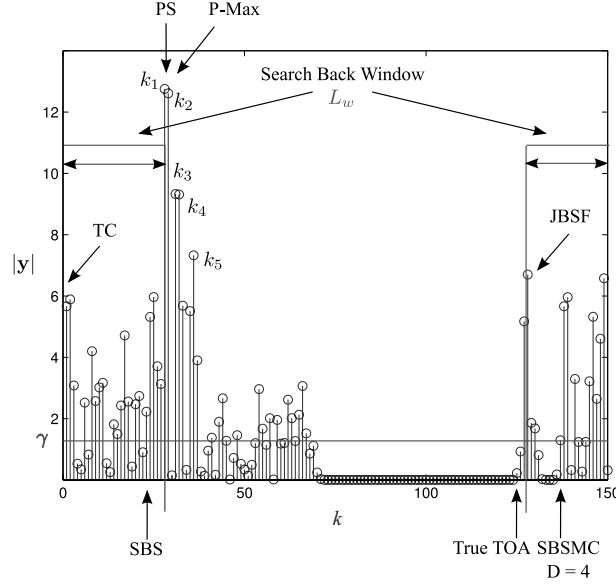
Figure 4.5: An example of applying different TOA strategies to a noiseless $|\mathbf{y}|$.

strongest path component, and then starts from the strongest path to search backward in a predefined window to look for the first threshold crossing. It has the problem that it may choose the index from a later arriving cluster instead of the first cluster. An example of the SBS method that has this problem is

Table 4.1: TOA Estimation Strategies

	Parameters	Algorithms
PS		$\hat{\tau} = T_b \underset{k \in \{1,\ldots,L_f\}}{\operatorname{argmax}} \ |[\mathbf{y}]_k| - T_b/2$
P-Max	P	$\hat{\tau} = T_b \min\left\{ k \in \{k_1,\ldots,k_P\} \right\} - T_b/2,$ $\begin{cases} k_1 = \underset{k \in \{1,\ldots,L_f\}}{\operatorname{argmax}} \ |[\mathbf{y}]_k| \\ k_i = \underset{k \in \{1,\ldots,L_f\} \wedge k \notin \{k_1,\ldots,k_{i-1}\}}{\operatorname{argmax}} \ |[\mathbf{y}]_k|, \\ i = 2,\ldots,P \end{cases}$
TC	γ	$\hat{\tau} = T_b \min\left\{ k \in \{1,\ldots,L_f\} \mid |[\mathbf{y}]_k| > \gamma \right\}$
JBSF	L_w, γ	$\hat{\tau} = T_b \min\left\{ k \in \{k_{\max} - L_w, \ldots, k_{\max}\} \mid |[\mathbf{y}]_k| > \gamma \right\}$
SBS	L_w, γ	$\hat{\tau} = T_b \max\left\{ k \in \{k_{\max},\ldots,k_{\max} - L_w\} \right.$ $\left. \mid (|[\mathbf{y}]_k| > \gamma) \wedge (|[\mathbf{y}]_{k-1}| < \gamma) \right\}$
SBSMC	L_w, γ, D	$\hat{\tau} = T_b \max\left\{ k \in \{k_{\max},\ldots,k_{\max} - L_w\} \right.$ $\mid (|[\mathbf{y}]_k| > \gamma) \wedge (|[\mathbf{y}]_{k-1}| < \gamma) \wedge \ldots$ $\left. \wedge (|[\mathbf{y}]_{k-D}| < \gamma) \right\}$

shown in Fig. 4.5. The serial backward search for multiple clusters (SBSMC) [82] works following the same principle as the SBS, but it additionally checks whether D consecutive samples before the candidate sample are all below the threshold or not. We can see the advantage of SBSMC over SBS in Fig. 4.5. Although it can partially solve the problem of the SBS method, it is still not favorable, because it needs more parameters related to the channel properties compared to other methods. Therefore, we employ the jump back and search forward (JBSF) method [77], which relies on two parameters, the threshold γ and the search back window length L_w. It first aligns with the strongest component, and jumps back to start from the index of $(k_{\max} - L_w)$ to find out the first $|\mathbf{y}|$ exceeding the threshold γ. In the example, the TOA estimate obtained by the JBSF method is the closest to the true TOA.

The optimal way to detect the leading edge would be a likelihood ratio test (LRT) [100] for multiple hypotheses. However, this requires full statistical knowledge of the channel, which is impractical. In the absence of channel information or lack of sufficient channel information, we could follow a heuristic approach and set the threshold to $\gamma_1 = \eta_t |[\mathbf{y}]_{k_{\max}}|$, where $0 \leq \eta_t \leq 1$ is the threshold ratio [101]. However, the optimal η_t would depend on the SNR in this scheme and there is no unique η_t that works well under all SNRs. On the other hand, we could set the threshold based on the stochastic properties of $|\mathbf{y}|$, when there is only noise. A threshold can be derived for a fixed probability of early false alarm P_{efa}, which indicates the event where we incorrectly select a noise sample before the true TOA as the first multipath component. If there is no absolute sample value exceeding the threshold in the predefined window, we use the result of the PS method $\hat{\tau} = T_b k_{max} - T_p/2$ as the estimated TOA. When there is only noise, $[\mathbf{y}]_k$ is an averaged AWGN sample with variance $\sigma^2 = \frac{1}{N_c(M-2)} N_0 B$. This means $|[\mathbf{y}]_k|$ follows a one-degree chi distribution (half-normal distribution).

$$\mathbf{y} \quad \overset{a}{\sim} \quad \mathcal{N}(0, \sigma^2 \mathbf{I}), \tag{4.18}$$

$$\frac{1}{\sigma}|\mathbf{y}| \quad \overset{a}{\sim} \quad \mathcal{X}_1, \tag{4.19}$$

$$P_{efa} \quad = \quad P(|\mathbf{y}| > \gamma_2; \text{noise only}) = 2Q(\frac{\gamma_2}{\sigma}), \tag{4.20}$$

$$\gamma_2 \quad = \quad \sigma Q^{-1}(\frac{P_{efa}}{2}), \tag{4.21}$$

where the function $Q(\cdot)$ is the right-tail probability function for a Gaussian distribution.

The length L_w of the backward search window should be large enough to recover the first path instead of deadlocking to the strongest one and depends on the delay between the strongest path and the first path, which could be as large as 60 ns for IEEE 802.15.4a CM1 [102] as shown in [88]. Due to the lack of channel knowledge, we choose L_w to be $\eta_l T_f$, where $0 \leq \eta_l \leq 1$ is the length ratio. We remark that T_f should also be long enough, not only to avoid IFI but also to

accommodate enough margin for the backward search window in order to prevent the window from including the channel tail. Therefore, T_f should be decided by the delay spread of the channel and the maximum delay between the strongest path and the first path together. For example, if we assume the channel length is 90 ns and the maximum delay between the first path and the strongest path is 60 ns, T_f should be at least 150 ns.

The complexity of the TOA estimator is negligible compared to the clock drift estimator. The peak selection finds the maximum value in $|\mathbf{y}|$. On the other hand, the JBSF method finds the maximum of $|\mathbf{y}|$, aligns the search back window with the maximum of $|\mathbf{y}|$, and then looks for the first $|[\mathbf{y}]_k|$ exceeding the threshold. The number of their operations is linear with L_f.

4.4 Theoretical performance limits

The main purpose of the CRB analysis here is to indicate the influence of the clock drift on the TOA estimation. In a multipath environment, TOA estimation consists of estimating the delay of the first multipath component. It can be regarded as part of the channel estimation. Although the parameters of other multipath components are nuisance parameters, we cannot get rid of them. Furthermore, we assume that the multipath components are independent and do not overlap with each other. Therefore, based on the received multipath signal, we would like to derive the CRB for $\boldsymbol{\theta} = [\Delta \ \tau_0 \ \tau_1 \ldots \tau_{L-1} \ \alpha_0 \ldots \alpha_{L-1}]^T$, which denotes the theoretical performance limits to jointly estimate the clock drift ratio and the multipath channel parameters. This CRB is different from the conventional CRBs [1,77,93,94], which only consider the multipath channel parameters. For instance, [93] and [94] calculate the CRBs for data-aided and non-data-aided ML channel estimation, respectively. We remark that the CRB derived here is not a favorable benchmark to compare with our practical TOA estimates, since our TOA estimation employs heuristic methods with low complexity. It is suboptimal, and its performance can never be beyond the resolution of the system. One the other hand, the CRB is derived based on the continuous waveform. Although the Ziv-Zakai bound (ZZB) is proposed in [103] to provide a tighter bound in the low SNR region, it has a similar behavior as the CRBs at the high SNR region. New bounds should be investigated to obtain useful benchmarks. It is left to future work.

Recall that τ_l and α_l are the delay and the amplitude of the lth received multipath component as we defined for (4.6) before. $\tau_0 = \tau$ is the TOA of interest. We assume that the multipath components do not overlap with each other. $\boldsymbol{\theta}$ is treated as deterministic unknown. The received signal $r(t)$ goes through the front-end filter to obtain

$$x(t) = \int_{-\infty}^{+\infty} r(z)g(t-z)dz, \tag{4.22}$$

where $g(t)$ is assumed to be the ideal bandpass filter, which includes the band of interest B as we defined before and $x(nT_{\text{sam}})$ defined in (4.9) is the nth sample of $x(t)$ using stroboscopic sampling. The output of the front-end filter $x(t)$ can be decomposed of a signal part and a noise part

$$x(t) \;=\; z(t) + n_g(t), \tag{4.23}$$

where $n_g(t)$ is the bandlimited AWGN with double sided power spectral density $N_0/2$ and bandwidth B, and

$$z(t) \;=\; \sum_{l=0}^{L-1} \alpha_l s_g(t - \tau_l, \Delta), \tag{4.24}$$

$$s_g(t, \Delta) \;=\; \sum_{i=0}^{N_c-1} c_i b_g(t - iN_f T_f(1 + \Delta), \Delta), \tag{4.25}$$

$$b_g(t, \Delta) \;=\; \sum_{j=0}^{N_f-1} \phi(t - jT_f(1 + \Delta)), \tag{4.26}$$

where $\phi(t) = \int_{-\infty}^{+\infty} p(\tau)g(t-\tau)d\tau$ is the one-pulse signal resulting from the channel, antenna and filter effects. We assume that $\phi(t)$ is known. Recalling $N_f = mM$ and $T_f = (mP + q)T_b$, we collect $mM(mP + q)N_c$ samples at rate $1/T_b$, while $M(mP+q)N_c$ samples at rate $1/T_{\text{sam}}$ from the above waveform. Since we exclude the first and last sample vector in each cluster to get rid of the code mismatch in the averaging process, we actually only use $(M - 2)(mP + q)N_c$ samples. We remark that if we use these stroboscopic samples instead of Nyquist rate samples to derive the CRBs, the new CRBs would be $\dfrac{mM}{M - 2}$ times the ones using Nyquist rate samples. The SNR penalty for the stroboscopic sampling compared to the Nyquist sampling is indicated by the reduced number of samples achieved from the same preamble.

The Fisher information matrix (FIM) $\mathbf{I}(\boldsymbol{\theta})$ is employed, with entries defined as:

$$[\mathbf{I}(\boldsymbol{\theta})]_{ij} \;=\; -E_{x(t);\boldsymbol{\theta}} \left[\frac{\partial^2}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{\theta}_j} \ln p\big(x(t); \boldsymbol{\theta}\big) \right], \tag{4.27}$$

where

$$\ln p\big(x(t); \boldsymbol{\theta}\big) = -\frac{1}{2\sigma^2} \int_0^{T_o} (x(t) - z(t))^2 dt \tag{4.28}$$

is the log-likelihood function without constant terms and T_o is the observation time. The derivation of the FIM and its inverse can be found in Appendix 4.A.

We can obtain the CRBs for Δ and τ based on the inverse of the FIM:

$$
\begin{aligned}
\mathrm{CRB}(\Delta) &= [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{1,1} && (4.29) \\
&= \frac{6}{\frac{\mathcal{E}_p}{N_0}\overline{F^2}N_f N_c T_f^2} \frac{1}{(N_f^2 N_c^2 - 1)\sum_{l=0}^{L-1}\alpha_l^2}, && (4.30) \\
\mathrm{CRB}(\tau) &= [\mathbf{I}^{-1}(\boldsymbol{\theta})]_{2,2} && (4.31) \\
&\approx \frac{1}{2\frac{\mathcal{E}_p}{N_0}\overline{F^2}N_f N_c}\left(\frac{1}{\alpha_0^2} + \frac{1}{\frac{1}{3}\sum_{l=0}^{L-1}\alpha_l^2}\right) && (4.32) \\
&\approx \frac{1}{2\frac{\mathcal{E}_p}{N_0}\overline{F^2}N_f N_c \alpha_0^2}. && (4.33)
\end{aligned}
$$

where $\mathcal{E}_p = \int_0^{T_o}\phi^2(t)dt$ is the pulse energy,

$$
\begin{aligned}
\overline{F^2} &= \int_0^{T_o}\left(\frac{d\phi(t)}{dt}\right)^2 dt / \int_0^{T_o}\phi^2(t)dt && (4.34) \\
&= \int_{-\infty}^{\infty}(2\pi F)^2|\Phi(F)|^2 dF / \int_{-\infty}^{\infty}|\Phi(F)|^2 dF, && (4.35)
\end{aligned}
$$

which is the mean square bandwidth of the pulse signal with $\Phi(F)$ the Fourier transform of $\phi(t)$. Note that $\overline{F^2}$ does not only depend on the bandwidth of the pulse signal $\phi(t)$, but also on its shape. A larger $N_c N_f$, which is the total number of pulses employed by the preamble, helps all of the estimates, because more frames and clusters enhance the SNR. A longer T_f facilitates the Δ estimation, since we can observe more drift with longer duration. $\mathrm{CRB}(\Delta)$ is independent of τ. Each multipath component contributes a part in the drift estimation. The more multipath energy we collect, the more accurate the drift estimate. $\mathrm{CRB}(\tau)$ is independent of Δ. The first approximation in the derivation of $\mathrm{CRB}(\tau)$ is bases on $N_f N_c \gg 1$. Comparing the two factors $1/\alpha_0^2$ and $1/(\frac{1}{3}\sum_{l=0}^{L-1}\alpha_l^2)$ in $\mathrm{CRB}(\tau)$, the former one is usually much larger than the later one. Therefore, the accuracy of the TOA estimation is mainly decided by the factor $1/\alpha_0^2$, which is the strength of the first path. The factor $1/(\frac{1}{3}\sum_{l=0}^{L-1}\alpha_l^2)$ introduced by the drift does not have much influence on the TOA estimation, since the drift is taken into account in the estimation approach.

If we assume that Δ or τ is known, we only have to estimate τ or Δ, respect-

ively. $\text{CRB}'(\Delta)$ or $\text{CRB}'(\tau)$ can be derived as

$$
\begin{aligned}
\text{CRB}'(\Delta) &= \frac{1}{[\mathbf{I}(\theta)]_{2,2}} && (4.36) \\
&= \frac{3}{\frac{\mathcal{E}_p}{N_0}\overline{F^2}N_f N_c T_f^2} \frac{1}{(2N_f^2 N_c^2 - 3N_f N_c + 1)\sum_{l=0}^{L-1}\alpha_l^2}, && (4.37) \\
\text{CRB}'(\tau) &= \frac{1}{[\mathbf{I}(\theta)]_{1,1}} && (4.38) \\
&= \frac{1}{2\frac{\mathcal{E}_p}{N_0}\overline{F^2}N_f N_c \alpha_0^2}. && (4.39)
\end{aligned}
$$

where $\text{CRB}'(\tau)$ is equal to the conventional CRB of τ in [1,77]. Because the estimation performance degrades when more parameters in a given system have to be estimated, we would have the relations as follows

$$
\begin{aligned}
\text{CRB}'(\Delta) &\leq \text{CRB}(\Delta), && (4.40) \\
\text{CRB}'(\tau) &\leq \text{CRB}(\tau). && (4.41)
\end{aligned}
$$

We will show examples of the CRBs in the next section.

4.5   Simulation results

The performance of TOA estimation is evaluated by simulations using the IEEE 802.15.4a channel model CM1 - indoor residential LOS [102]. The channel impulse responses are truncated to 90 ns to avoid IFI in order to simplify the simulation. In practice, we can increase the frame period to avoid IFI. One hundred channel realizations are generated, and we randomly choose one for each Monte Carlo run. To speed up the simulations, we generate the output of the cluster averaging process directly, since the maximum drift of T_b within a cluster is smaller than the pulse width. The averaging process would not suffer from severe pulse mismatch as we mentioned before. E_c/N_o defines the cluster energy to noise variance ratio. The number of clusters N_c is 68. Further, we randomly select the drift ratio Δ among $\{-67.2$ ppm, -50.4 ppm, -33.6 ppm, -16.8 ppm, 0 ppm, 16.8 ppm, 33.6 ppm, 50.4 ppm, 67.2 ppm$\}$ in each run. The frame period T_f is 150 ns, which not only depends on the channel length, but also on the maximum delay between the first path and the strongest path. Since the maximum delay between the first path and the strongest path is 60 ns for CM1 [88], the frame period is set as the sum of the channel length and the maximum delay. The stroboscopic sampling period T_{sam} is chosen to be 7 ns. The targeted resolution T_b is 1 ns. Based on (4.2), we obtain $P = 21$, $m = 7$ and $n = 3$. The maximum M is 12 according to (4.4). The number of frames in one cluster $N_f = mM = 84$. Therefore, the processing gain of the cluster averaging process is $M - 2 = 10$, or approximately
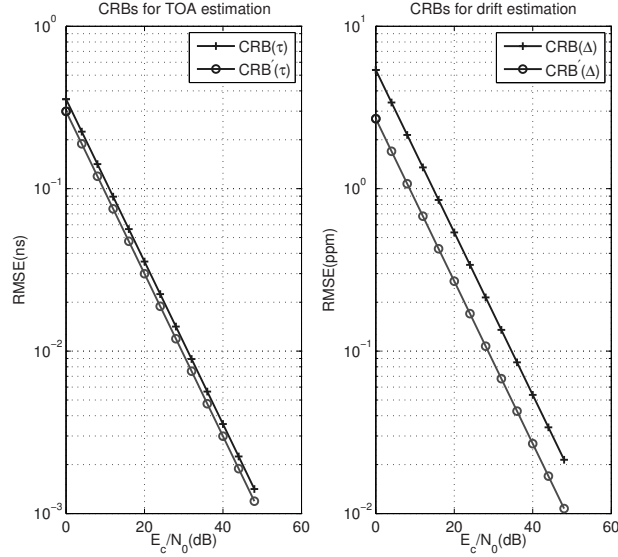
Figure 4.6: CRBs for TOA and drift estimation

10 dB. The total processing gain is $N_c(M-2) = 680$, or approximately 28 dB. The second derivative of a Gaussian pulse is employed as the transmitted pulse, and the pulse width is truncated to 4 ns, which includes the main lobe and two side lobes. Its bandwidth is approximately 500 MHz. The pulse is also used as the impulse response of the front-end filter at the receiver. Moreover, the timing offset $\tau \in [0, T_f)$ is randomly generated in each run. The threshold ratio is $\eta_t \in \{0.1, 0.2, \ldots, 0.8\}$. The length ratio of the backward search window is $\eta_l \in \{0.1, 0.2, \ldots, 0.7\}$, which leads to $L_w \in \{10 \text{ ns}, 20 \text{ ns}, \ldots, 70 \text{ ns}\}$.

We first show the CRBs under above simulation conditions. The CRBs are much lower than the performance of the PS method or the JBSF method shown later. It is because that for the CRB derivation, we use the assumption that the multipath components are independent and do not overlap with each other. We also have the prior knowledge of the number of the multipaths. But in the simulations, we only use a signal with a bandwidth of 500 MHz, the multipaths may not be resolvable. Therefore, the suboptimal TOA estimation methods have performance gaps compared to the CRBs. As we can see in Fig. 4.6, $\text{CRB}(\tau)$ is inferior to $\text{CRB}'(\tau)$. The difference is very small, which indicates that the influence of the drift is very small in the TOA estimation, when it is taken into consideration. $\text{CRB}'(\Delta)$ is lower than $\text{CRB}(\Delta)$ since we assume that τ is known to derive the former one.

Secondly, we evaluate the performance of TOA estimation using the PS method,
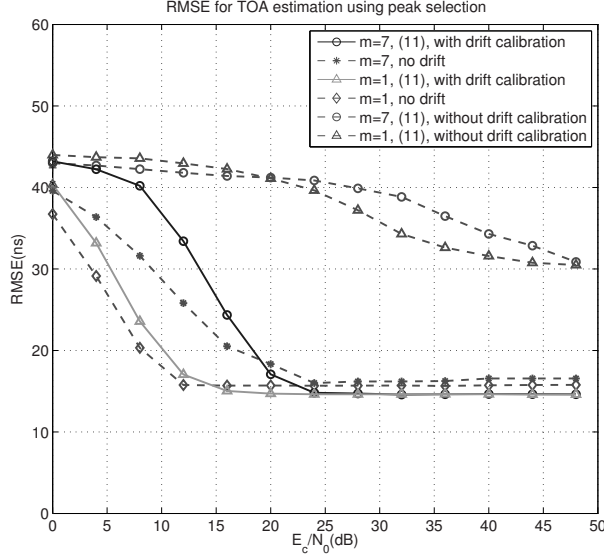
Figure 4.7: RMSE of TOA estimation using peak selection.

which consists of choosing the strongest component. The root mean square error (RMSE) of $\hat{\tau}_p$ vs. E_c/N_o is illustrated in Fig. 4.7. There are large performance gaps between the cases with drift calibration (solid lines) and the cases without drift calibration (dashed lines with $\circ$ or $\triangle$ markers). The drift causes serious problems to TOA estimation as indicated by the high error floor for the curves without calibration. The Nyquist sampling system, where $m = 1$ and $n = 0$, is used as a reference. The same preamble is used for the Nyquist sampling system, the stroboscopic sampling system and the CRB derivation. No matter how large m, the RMSE of TOA estimation with drift calibration always converges to the same error floor (solid lines). Nyquist sampling gains about 8 dB over strobo-scopic sampling ($m = 7$) due to the sampling-rate gain m. For both systems, we also show the performance of the ideal case without any drift. They show similar error floors as the ones with drift calibration, which proves that the drift is not the reason for the error floor. However, due to the fact that we employ a signal with a bandwidth of 500 MHz, we may not be able to resolve fine multipaths. Because of inter-pulse interference (IPI), the suboptimal PS method always chooses the strongest resolvable signal component instead of the strongest single path, which causes a high error floor. We may improve the performance of the PS method by employing a signal with a larger bandwidth. Furthermore, the first path is not always the strongest one in CM1, which is also indicated by the error floor.

We also investigate the performance of TOA estimation using the JBSF method,
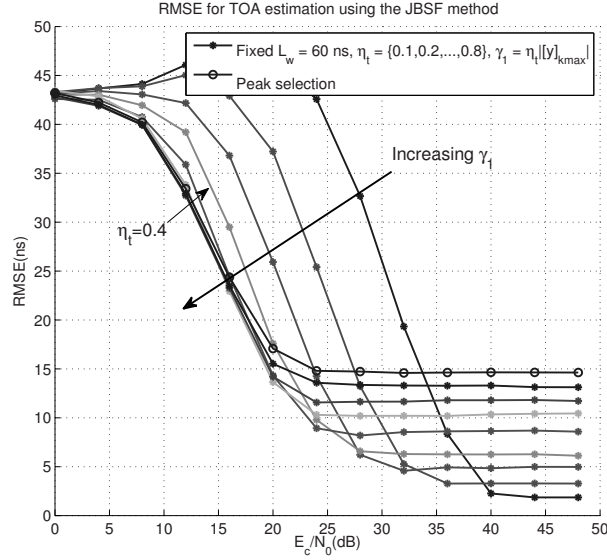
Figure 4.8: RMSE of TOA estimation using the JBSF method for a fixed window length $L_w = 60$ ns and different thresholds $\gamma_1 = \eta_t |[\mathbf{y}]_{k_{max}}|$, $(m = 7)$.

which achieves much better accuracies than the PS method. Fig. 4.8 shows the RMSE of $\hat{\tau}$ for a fixed window length $L_w = 60$ ns and different thresholds $\gamma_1 = \eta_t |[\mathbf{y}]_{k_{max}}|$. There is no single threshold which could outperform the others over the whole SNR range. A smaller threshold can achieve a better accuracy at high SNR, but also performs worse at low SNR, which is consistent with the conclusions in [101]. The error floor of peak selection is higher than any error floor encountered by the JBSF method. According to Fig. 4.8, the performance of $\eta_t = 0.4$ yields a good tradeoff over the whole SNR range. It only has a 2dB performance loss in the SNR range of $0 - 25$dB compared to larger thresholds, and its error floor is still close to the low error floors obtained by smaller thresholds. Furthermore, the performance of the JBSF method with a fixed threshold $\gamma_1 = 0.4|[\mathbf{y}]_{k_{max}}|$ under different backward search windows is shown in Fig. 4.9. The larger the window length, the smaller the error floor. The performance differences in the low SNR range are smaller than the ones with a fixed window length and various thresholds γ_1. There is no more performance improvement, when the window length is larger than 60 ns. Since we know that 60 ns is the best window length according to the maximum delay between the first path and the strongest path, a window length larger than 60ns may introduce the problem of regarding the channel tail as the leading edge.

When we observe the performance of the JBSF method using different thresholds
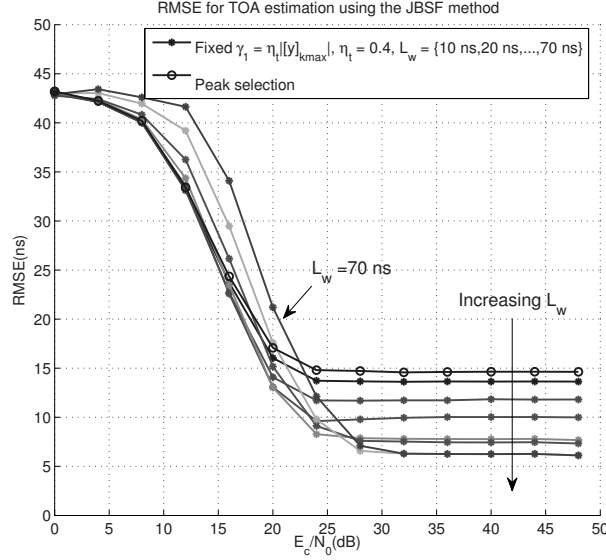
Figure 4.9: RMSE of TOA estimation using the JBSF method for a fixed threshold $\gamma_1 = \eta_t |[\mathbf{y}]_{k_{max}}|, \eta_t = 0.4$ and different window lengths, $(m = 7)$.

$\gamma_2 = \sigma Q^{-1}(P_{efa}/2)$ for a fixed window length $L_w = 60$ ns in Fig. 4.10, they all have the same performance in the SNR range of $0 - 25$dB as the PS method. This is because in the low SNR range, the received signal is overwhelmed by noise. The chance of a sample value exceeding the threshold is low. As we use the result of peak selection when there is no threshold crossing, the performance of the JBSF method using γ_2 in the low SNR range is the same as the PS method. In the high SNR range, the larger the P_{efa}, the better the TOA estimate. The PS method still has the highest error floor. Further, we use a fixed threshold related to $P_{efa} = 0.1$ to test the estimation performance under different window lengths as shown in Fig. 4.11. The larger the window length, the lower the error floor. The exception happens when $L_w = 70$ ns. Then the search window is too long, and includes the channel tail, which causes an estimation error. This observation is consistent with Fig. 4.9.

In Fig. 4.12, we finally compare the performance of the JBSF method for $\gamma_1 = \eta_t |[\mathbf{y}]_{k_{max}}|, \eta_t = 0.4$ and $\gamma_2 = \sigma Q^{-1}(P_{efa}/2), P_{efa} = 0.1$ using a fixed window length $L_w = 60$ ns. There are still large performance gaps between the methods with drift calibration and the ones without. In general, the PS method has a higher error floor than the JBSF method. The JBSF method using γ_2 could converge to a lower error floor than the JBSF method using γ_1, but it has a worse performance in the middle SNR range of $20 - 30$ dB for $m = 7$ ($11 - 28$dB for
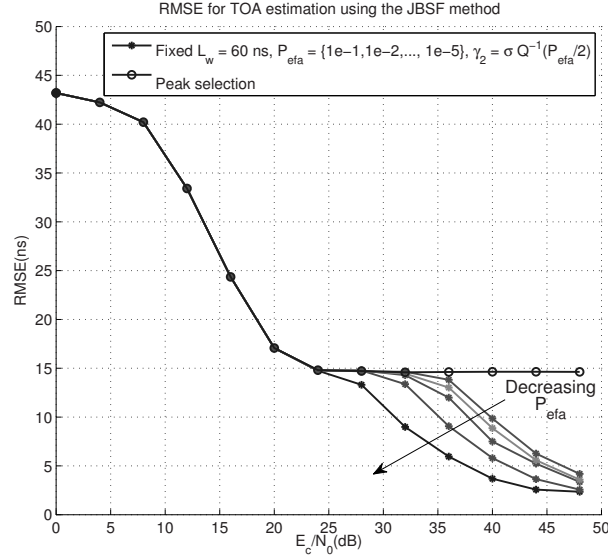
Figure 4.10: RMSE of TOA estimation using the JBSF method for a fixed window length $L_w = 60$ ns and different threshold $\gamma_2 = \sigma Q^{-1}(P_{efa}/2)$, ($m = 7$).

$m = 1$). Based on the above observations, we could employ γ_2 in the low and middle SNR range and γ_1 at high SNR.

4.6 Conclusions

In this chapter, we have applied stroboscopic sampling for an IR-UWB system to achieve accurate TOA estimation with a low sampling rate. Due to the long preamble required by stroboscopic sampling, the clock drift is one of the main error sources in TOA estimation. Hence, we include the drift into our system model and obtain an drift estimate by an exhaustive search. Further, we apply peak selection and the JBSF method to estimate the TOA using the averaged data samples corrected for the drift. Simulation results confirm that the drift calibration dramatically reduces the TOA estimation errors due to the drift, and stroboscopic sampling can achieve the same estimation resolution as Nyquist sampling. We have proposed a practical low sampling rate solution for TOA estimation using UWB signals.
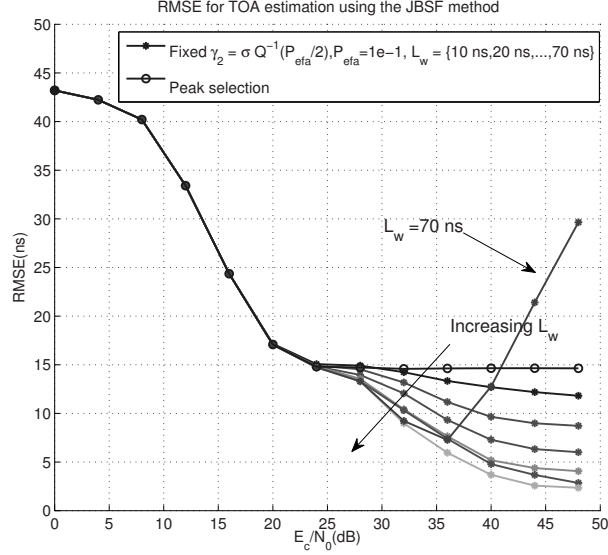
Figure 4.11: RMSE of TOA estimation using the JBSF method for a fixed threshold $\gamma_2 = \sigma Q^{-1}(P_{efa}/2)$, $P_{efa} = 0.1$ and different window lengths, $(m = 7)$.

4.A   CRB derivation

We assume that the multipath components do not overlap with each other. Then the derivatives of $z(t)$ w.r.t. Δ, τ_l and α_l are given respectively by

$$\frac{\partial z(t)}{\partial \Delta} = \sum_{l=0}^{L-1} \alpha_l \frac{\partial s_g(t - \tau_l, \Delta)}{\partial \Delta}, \tag{4.42}$$

$$\frac{\partial z(t)}{\partial \tau_l} = -\alpha_l \frac{\partial s_g(t - \tau_l, \Delta)}{\partial (t - \tau_l)}, \tag{4.43}$$

$$\frac{\partial z(t)}{\partial \alpha_l} = s_g(t - \tau_l, \Delta). \tag{4.44}$$

Next, we calculate each entry of the FIM. We assume that there is no IFI than T_f and the observation window T_o includes the whole received waveform, it leads
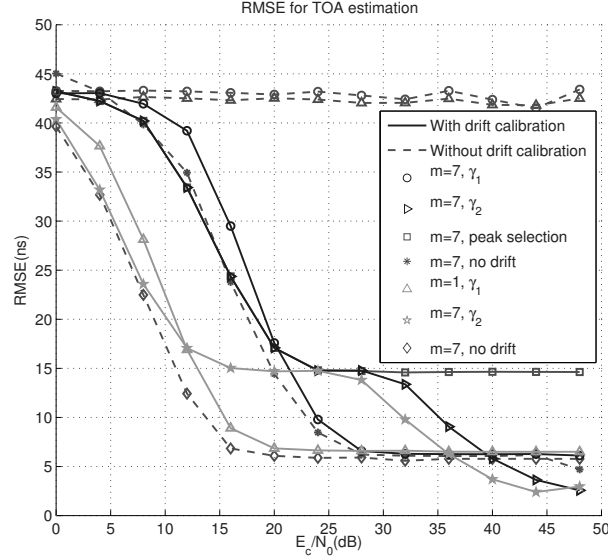
Figure 4.12: RMSE of TOA estimation using the JBSF method, $L_w = 60$ ns, $\gamma_1 = \eta_t |[\mathbf{y}]_{k_{max}}|, \eta_t = 0.4$ and $\gamma_2 = \sigma Q^{-1}(P_{efa}/2), P_{efa} = 0.1$.

us to

$$
\begin{aligned}
E_{x(t);\boldsymbol{\theta}}\left[\frac{\partial^2 \mathrm{ln}p(x(t);\boldsymbol{\theta})}{\partial \Delta^2}\right] &= -\frac{2}{N_0}\int_0^{T_o}\left(\frac{\partial z(t)}{\partial \Delta}\right)^2 dt \\
&= -\frac{\mathcal{E}_p}{3N_0}\overline{F^2}T_f^2 N_f N_c(2N_f^2 N_c^2 - 3N_f N_c + 1)\sum_{l=0}^{L-1}\alpha_l^2.
\end{aligned}
\tag{4.45}
$$

$$
\begin{aligned}
E_{x(t);\boldsymbol{\theta}}\left[\frac{\partial^2 \mathrm{ln}p(x(t);\boldsymbol{\theta})}{\partial \tau_l \partial \Delta}\right] &= -\frac{1}{\sigma^2 T_s}\int_0^{T_o}\frac{\partial z(t)}{\partial \tau_l}\frac{\partial z(t)}{\partial \Delta}dt \\
&= -\frac{\mathcal{E}_p}{N_0}\overline{F^2}\alpha_l^2 T_f N_f N_c(N_f N_c - 1).
\end{aligned}
\tag{4.46}
$$

$$
\begin{aligned}
E_{x(t);\boldsymbol{\theta}}\left[\frac{\partial^2 \mathrm{ln}p(x(t);\boldsymbol{\theta})}{\partial \alpha_l \partial \Delta}\right] &= -\frac{2}{N_0}\int_0^{T_o}\frac{\partial z(t)}{\partial \alpha_l}\frac{\partial z(t)}{\partial \Delta}dt \\
&= \frac{2}{N_0}\sum_{i=0}^{N_c-1}\sum_{j=0}^{N_f-1}(iN_f + j)T_f\alpha_l \\
&\quad \times \int_{-(iN_f+j)T_f-\tau_l}^{T_o-(iN_f+j)T_f-\tau_l}\left(\phi(t)\frac{d\phi(t)}{dt}\right)dt,
\end{aligned}
\tag{4.47}
$$

where $\int_0^{T_o} \phi(t)\frac{\phi(t)}{dt}dt = \int_0^{T_o} \phi(t)d\phi(t) = \frac{\phi^2(t)}{2}\Big|_0^{T_o}$. We assume $\phi(0) = \phi(T_o) = 0$.
It leads us to $\int_0^{T_o} \phi(t)\frac{\phi(t)}{dt}dt = 0$. The above equation equals to zero.

$$
\begin{aligned}
E_{x(t);\boldsymbol{\theta}}\left[\frac{\partial^2 \ln p(x(t);\boldsymbol{\theta})}{\partial \alpha_l \partial \tau_v}\right] &= -\frac{2}{N_0}\int_0^{T_o}\frac{\partial z(t)}{\partial \alpha_l}\frac{\partial z(t)}{\partial \tau_v}dt \\
&= 0.
\end{aligned} \tag{4.48}
$$

$$
\begin{aligned}
E_{x(t);\boldsymbol{\theta}}\left[\frac{\partial^2 \ln p(x(t);\boldsymbol{\theta})}{\partial \tau_l \partial \tau_v}\right] &= -\frac{2}{N_0}\int_0^{T_o}\frac{\partial z(t)}{\partial \tau_l}\frac{\partial z(t)}{\partial \tau_v}dt \\
&= \begin{cases} 0 & l \neq v \\ -\frac{2\mathcal{E}_p}{N_0}\overline{F^2}\alpha_l^2 N_f N_c & l = v \end{cases}.
\end{aligned} \tag{4.49}
$$

$$
\begin{aligned}
E_{x(t);\boldsymbol{\theta}}\left[\frac{\partial^2 \ln p(x(t);\boldsymbol{\theta})}{\partial \alpha_l \partial \alpha_v}\right] &= -\frac{2}{N_0}\int_0^{T_o}\frac{\partial z(t)}{\partial \alpha_l}\frac{\partial z(t)}{\partial \alpha_v}dt \\
&= \begin{cases} 0 & l \neq v \\ -\frac{2\mathcal{E}_p}{N_0} N_f N_c & l = v \end{cases}.
\end{aligned} \tag{4.50}
$$

Filling in all the entries, we calculate the inverse of the FIM based on the block matrix inversion as

$$
\mathbf{I}^{-1}(\boldsymbol{\theta}) = \frac{1}{2\frac{\mathcal{E}_p}{N_0}\overline{F^2}N_f N_c}\begin{bmatrix} \frac{1}{d} & -\frac{1}{d}\mathbf{b}^T\mathbf{A}^{-1} \\ -\frac{1}{d}\mathbf{A}^{-1}\mathbf{b} & (\mathbf{A} - \frac{1}{c}\mathbf{b}\mathbf{b}^T)^{-1} \end{bmatrix}, \tag{4.51}
$$

where

$$
\mathbf{b} = \frac{T_f}{2}(N_f N_c - 1)[\alpha_0^2 \quad \alpha_1^2 \quad \dots \quad \alpha_{L-1}^2 \quad \mathbf{0}_L^T]^T, \tag{4.52}
$$

$$
c = \frac{T_f^2}{6}(2N_f^2 N_c^2 - 3N_f N_c + 1)\sum_{l=0}^{L-1}\alpha_l^2, \tag{4.53}
$$

$$
\mathbf{A} = \text{diag}([\alpha_0^2 \quad \alpha_1^2 \quad \dots \quad \alpha_{L-1}^2 \quad \frac{1}{\overline{F^2}}\mathbf{1}_L^T]), \tag{4.54}
$$

$$
\begin{aligned}
d &= c - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b} \\
&= \frac{T_f^2}{12}(N_f^2 N_c^2 - 1)\sum_{l=0}^{L-1}\alpha_l^2.
\end{aligned} \tag{4.55}
$$

According to the Woodbury identity [104], the bottom right entry can be rewritten

as

$$\left(\mathbf{A} - \frac{1}{c}\mathbf{b}\mathbf{b}^T\right)^{-1} = \mathbf{A}^{-1} + \frac{\mathbf{A}^{-1}\mathbf{b}\mathbf{b}^T\mathbf{A}^{-1}}{c - \mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}} \tag{4.56}$$

$$= \mathbf{A}^{-1} + \frac{1}{\frac{1}{3}\left(1 + \frac{2}{N_f N_c - 1}\right)\sum_{l=0}^{L-1}\alpha_l^2} \left[\begin{array}{cc} \mathbf{1}_L\mathbf{1}_L^T & \mathbf{0}_L\mathbf{0}_L^T \\ \mathbf{0}_L\mathbf{0}_L^T & \mathbf{0}_L\mathbf{0}_L^T \end{array}\right].$$

Part III

Localization and Tracking
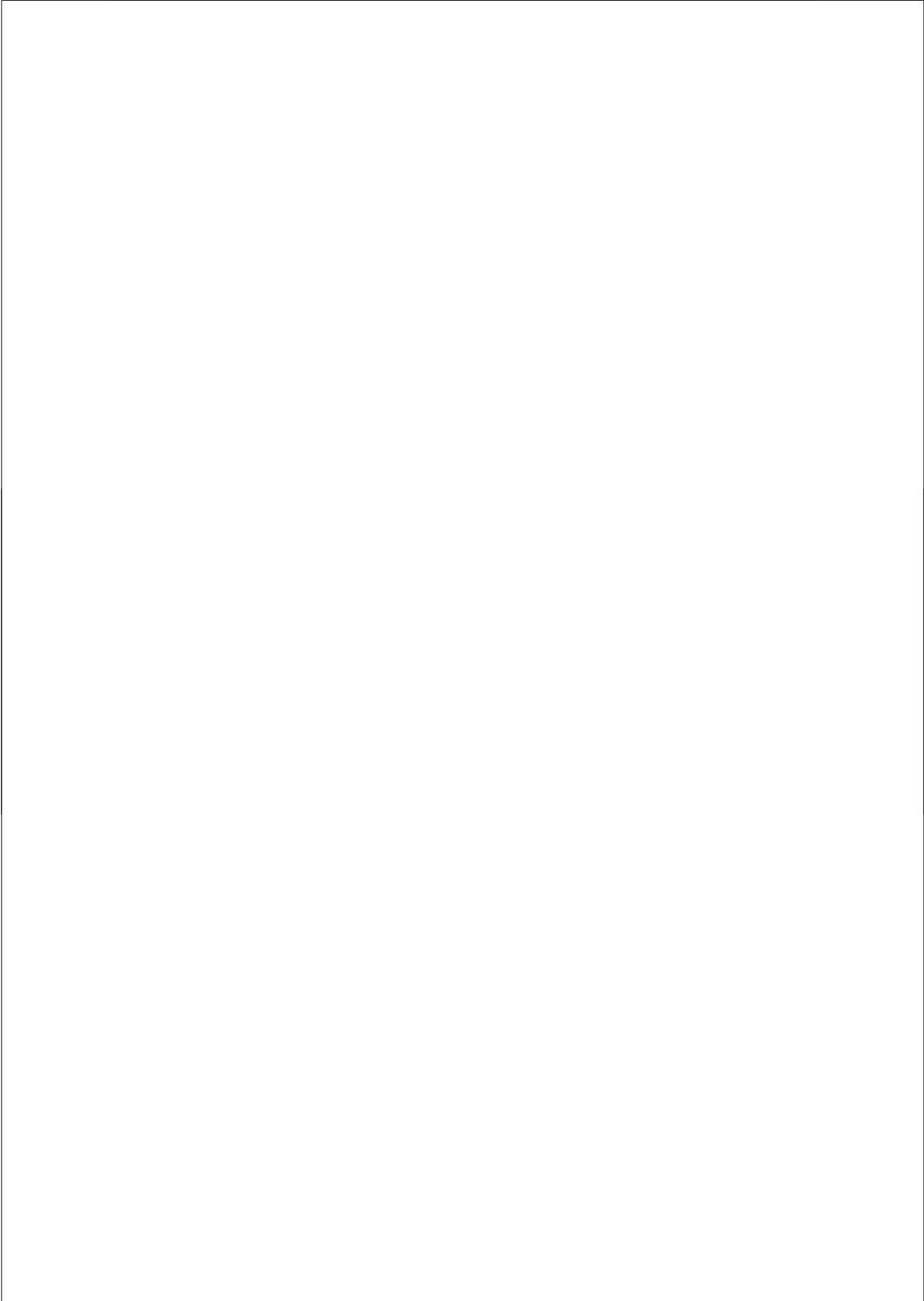
In this part, we focus on localization and tracking for WSNs, which have a wide range of applications, such as target tracking, surveillance, environment monitoring, geographical routing and smart home, and also impose a number of particular challenges on accuracy, complexity, and robustness for localization and tracking. We remark that the approaches proposed in this part can not only be applied to WSNs, but also be adopted to other localization systems, such as the global positioning system (GPS) [3], acoustic systems [6, 7], cellular networks [8], etc. In general, the range-based methods follow two steps [1, 2, 9]: firstly measure the metrics bearing location information, the so-called ranging or bearing, and secondly estimate the positions based on those metrics, the so-called location information fusion. There are mainly four metrics: time-of-arrival (TOA) or time-of-flight (TOF) [11], time-difference-of-arrival (TDOA) [6, 12], angle-of-arrival (AOA) [13], and received signal strength (RSS) [14]. The ranging methods using RSS can be implemented by energy detectors, but they can only achieve a coarse resolution. Antenna arrays are required for AOA-based methods, which encumbers their popularity. On the other hand, the high accuracy and potentially low cost implementation make TOA or TDOA based on ultra-wideband impulse radios (UWB-IRs) a promising ranging method [1].

This part consists of four chapters: Chapter 5-8. As we mentioned in Chapter 1, localization methods can be categorized as singular or aggregate approaches based on whether there are enough anchors (nodes with known positions) to assist the localization for each target sensor node. In Chapter 5, we assume there are no anchors, and propose an extension of the classical multidimensional scaling (MDS) algorithm to estimate the positions of all the nodes given partial pairwise distance measurements up to a translation, rotation, and reflection, as the classical MDS can only deal with fully connected networks. Moreover, the proposed algorithm can be categorized as an aggregate localization method.

Note that in Chapter 5 we do not care about how to obtain the distance measurements. Hence, we intentionally isolate the ranging from the ranging fusion in order to simplify the localization problem. However, this strict separation between the ranging and the ranging fusion burdens the first one and relieves the second one. It is difficult to obtain accurate range measurements independently. Therefore, we try to balance these two steps. The cross-layer philosophy inspires us to explicitly model ranging errors and handle them in the ranging fusion step. It provides a totally new look at the localization problem. From Chapter 6 to Chapter 8, we employ TOA (or TDOA) as range measurements, and thus clock synchronization becomes essential for ranging to achieve accurate localization. Based on the cross-layer philosophy, we take clock parameters into account to develop the localization algorithms. In these chapters, we consider M anchor nodes and one target node, and would like to estimate or track the position of the target node. All the nodes are again distributed in an l-dimensional space. The coordinates of the anchor nodes are known and defined as $\mathbf{X}_a = [\mathbf{x}_1, \ \mathbf{x}_2, \ \ldots, \ \mathbf{x}_M]$, where the vector $\mathbf{x}_i = [x_{1,i}, \ x_{2,i}, \ \ldots, \ x_{l,i}]^T$ of length l indicates the known coordinates of the ith anchor node. We employ a vector $\mathbf{x}$ of length l to denote the unknown

Synchronization level w.r.t. anchors and Chapters	Clock model
Synchronous Chapter 6	$C_i(t) = t + \theta, \ \ i = 1, 2, \ldots, M,$ where θ is the unknown common clock offset.
Quasi-synchronous Chapter 7 and 8	$C_i(t) = t + \theta_i, \ \ i = 1, 2, \ldots, M,$ where θ_i is the unknown clock offset.
Fully asynchronous Chapter 7	$C_i(t) = \alpha_i t + \theta_i, \ \ i = 1, 2, \ldots, M,$ where α_i is the unknown clock skew.

Table 4.2: Clock models for anchor nodes.

coordinates of the target node. The localization and tracking approaches proposed in these chapters are singular methods. We remark that these methods can easily be extended for multiple target nodes. Therefore, we can locate the target nodes in a sequential way for a large scale WSN [27]. The target nodes that have enough anchors are localized first. Then, the located target nodes can be viewed as new anchors that can facilitate the localization of other target nodes. As a result, the multiple-anchors-one-target scenario discussed from Chapter 6 to Chapter 8 is of practical interest. Furthermore, we consider three different kinds of networks with different synchronization levels w.r.t. to the anchor clock. Note that these synchronization requirements are related to the anchors, and no synchronization requirements are upon the target node, which indicates that the clock of the target node can run freely. Strictly speaking, they are all asynchronous networks. Consequently, the target clock $C_s(t)$ can be modeled as [29]

$$C_s(t) \quad = \quad \alpha_s t + \theta_s, \tag{4.57}$$

where α_s and θ_s denote the unknown clock skew and clock offset of the target node clock. We summarize all the anchor clock models for Chapter 6 - 8 in Table 4.2.

Based on the same cross-layer philosophy, different chapters focus on different aspects. Chapter 6 investigates reference-free localization estimators based on TOA measurements with the help of synchronous anchors. Instead of using TDOAs to allow an asynchronous target, we model the range error due to the asynchronous effect as a common bias, and propose various reference-free localization estimators. The reference-free LS estimators that are different from the reference-based ones do not suffer from a poor reference selection. Furthermore, we expose the relations among existing reference-based localization estimators and the proposed ones in order to clarify some confusions that still persist in recent literature. In Chapter 7, we propose robust TOA-based localization strategies to locate a target node in quasi-synchronous and fully asynchronous networks, respectively. Time-based localization is tightly coupled with clock synchronization. Thus, we explicitly take the clock parameters into account in our data model,

and further we consider the reliability of timestamps. A novel ranging protocol is developed, namely asymmetric trip ranging (ATR), to reduce the communication load and explore the broadcast property of WSNs. Regardless of the reliability of the timestamp report from the target node, several closed-form least-squares (LS) estimators are derived. As a result, we counter the uncertainties caused by the target node by ignoring the timestamps from this node. Since we have already developed approaches to linearize the measurement models in the previous chapters, a Kalman filter (KF) based tracking approach follows naturally in Chapter 8 to track a target node in quasi-synchronous networks. We again employ the ATR protocol, and inherit the linearization methods from the previous chapters to derive a linear measurement model from the TOA measurements. Thus, the proposed KF has the advantages of a zero modeling error and a low computational complexity compared to the extended KF.

Extending the Classical MDS

5.1 Introduction

Node localization from pairwise Euclidean distance measurements has become a fundamental research topic with the growing interest in wireless sensor networks [2]. The classical MDS algorithm [106] transforms the problem into an eigenvalue problem of a so-called projected distance matrix which solely depends on the distances. This matrix has rank two (for a planar geometry) and its first two eigenvectors provide the coordinates of all nodes up to a translation, rotation, and reflection [107–111]. Applying the MDS algorithm requires a fully connected sensor network, i.e., given a sensor network with N nodes, we need $N(N-1)/2$ distance measurements.

However, in practice, only a limited number of distance measurements are given due to communication limits imposed on the nodes (e.g., battery constraints). For simplicity, assume we have an ad-hoc sensor network with N nodes where the positions of all the nodes are unknown. Only M nodes (in practice, $M \ll N$) in the network can communicate with all the other nodes. Such nodes can be anchor nodes which usually have less stringent communication constraints. We assume that the other nodes $N - M$ cannot communicate with each other due to communication limits, but communicate with the first nodes. One possible solution to this setup is the "scaling by majorizing a complicated function" (SMACOF) algorithm, which is based on iteratively minimizing a global stress

---

cost function [107, 112] composed of a weighted least squares (LS) part and possibly a penalty term that includes prior information about node positions. This approach involves a highly nonlinear cost function which requires many initial guesses of the node positions to ensure convergence to the global minimum. As a result, suboptimal solutions have been investigated, which can possibly be used as an initial point of this algorithm. Examples are MDS-MAP [110], and SVD-Reconstruct [113], which both are based on first completing the missing entries in the distance matrix and then applying the classical MDS, or Nyström's algorithm [114] which is used to reduce the complexity of the singular value decomposition step involved in the classical MDS algorithm. With Nyström's algorithm the positions of the first group are estimated using the classical MDS method, while the positions of the second group are estimated using the LS method based on the former results and the mutual measurements between the two groups [10, eqs. (8)-(9)].

Herein, we develop a localization approach which extends the classical MDS to the current setup. We localize the nodes of the first group by exploiting their distance measurements to the disconnected nodes. At the same time, the second group's positions are obtained up to a translation relative to the former nodes. If recovering this translation is also of interest, then a possible additional step can be implemented. We examine our results with Monte-Carlo simulations by evaluating the normalized root mean square error (RMSE) between the true inter-node distances and their estimates. We compare our two-step approach with the MDS-MAP [110], the SVD-Reconstruct [113], and Nyström's algorithm [114]. Simulation results show that our proposed approach outperforms these previous solutions.

5.2   Problem formulation

Consider N sensor nodes randomly distributed in a two-dimensional plane. Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N] \in \mathcal{R}^{2 \times N}$ consist of all the node positions, where $\mathbf{x}_n \in \mathcal{R}^{2 \times 1}$, $n = 1, 2, \ldots, N$, is the position of the nth node. We assume that we have two groups of nodes: 1) M nodes which are fully connected; 2) $N - M$ nodes which are fully disconnected (i.e., each of the nodes in the second group does not communicate with the other nodes in this group, but communicates with each node in the first group.) Let $\mathbf{X}_1 = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_M] \in \mathcal{R}^{2 \times M}$ and $\mathbf{X}_2 = [\mathbf{x}_{M+1}, \ldots, \mathbf{x}_N] \in \mathcal{R}^{2 \times (N-M)}$ represent the set of node positions of the first group and the second group, respectively. The distance measurement between the ith and jth nodes (given they are connected) is $r_{i,j} = d_{i,j} + e_{i,j}$, where $d_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\|$, and $e_{i,j} \sim \mathcal{N}(0, \sigma^2)$ is the uncorrelated additive noise, where σ^2 is a known noise variance. The problem discussed herein is briefly stated as follows: Given the available pairwise measurements $\{r_{i,j}\}$, determine the positions of the nodes (up to rotation, reflection and translation). In other words, our goal is to reconstruct the constellation of the sensor network given partial pairwise node measurements.

5.3   Review of the classical MDS

Let us first review the classical MDS [112] in the noiseless case. For a fully connected network, we collect all the true distances in $\mathbf{D}$, where $[\mathbf{D}]_{i,j} = [\mathbf{D}]_{j,i} = d_{i,j}$. By expanding $d_{i,j}^2 = \|\mathbf{x}_i - \mathbf{x}_j\|^2 = \|\mathbf{x}_i\|^2 + \|\mathbf{x}_j\|^2 - 2\mathbf{x}_i^T\mathbf{x}_j$, we can model $\mathbf{D}$ as

$$\mathbf{D} \odot \mathbf{D} \;\; = \;\; \boldsymbol{\psi}\mathbf{1}_N^T - 2\mathbf{X}^T\mathbf{X} + \mathbf{1}_N\boldsymbol{\psi}^T, \qquad (5.1)$$

where $\boldsymbol{\psi} = \left[\|\mathbf{x}_1\|^2, \ldots, \|\mathbf{x}_N\|^2\right]$, and $\mathbf{1}_n \in \mathcal{R}^{n \times 1}$ is a vector with all elements equal to one. Based on (5.1), we would like to estimate $\mathbf{X}$. We employ an orthogonal projection $\mathbf{P}$ onto the orthogonal complement of $\mathbf{1}_N$, which is given by $\mathbf{P} = \mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T$. Since $\mathbf{P}\mathbf{1}_N = \mathbf{0}_N$, $\mathbf{P}$ can be used to eliminate the terms $\boldsymbol{\psi}\mathbf{1}_N^T$ and $\mathbf{1}_N\boldsymbol{\psi}^T$ at the right hand side of (5.1). Therefore, pre- and post- multiplying $\mathbf{P}$ to both sides of (5.1), we obtain

$$\mathbf{P}(\mathbf{D} \odot \mathbf{D})\mathbf{P} \;\; = \;\; -2\mathbf{P}\mathbf{X}^T\mathbf{X}\mathbf{P}, \qquad (5.2)$$

where $\mathbf{P}\mathbf{X}$ is a shifted version of $\mathbf{X}$, and the gravity of $\mathbf{P}\mathbf{X}$ is in the origin. Note that $\mathbf{P}\mathbf{X}$ is a $N \times 2$ "tall" matrix, which can be recovered up to a rotation and reflection. Then, we can recover $\mathbf{X}$ up to a translation, rotation and reflection by the eigenvalue decomposition of $-\frac{1}{2}\mathbf{P}(\mathbf{D} \odot \mathbf{D})\mathbf{P}$, and we achieve

$$\mathbf{X} \;\; = \;\; \mathrm{diag}(\lambda_1^{1/2}, \lambda_2^{1/2})[\mathbf{U}]_{:,1:2}^T, \qquad (5.3)$$

where $-\frac{1}{2}\mathbf{P}(\mathbf{D} \odot \mathbf{D})\mathbf{P} = \mathbf{U}\mathrm{diag}(\lambda_1, \ldots, \lambda_N)\mathbf{U}^T$ is the eigenvalue decomposition of $-\frac{1}{2}\mathbf{P}(\mathbf{D} \odot \mathbf{D})\mathbf{P}$ with $\mathbf{U}$ collecting the eigenvectors and λ_i the ith eigenvalue.

5.4   Possible position projections

As the classical MDS requires a fully connected network, it cannot be directly applied to our partially connected network in Section 5.2. Thus, we tailor it for our problem. Let us again consider the noiseless case ($r_{i,j} = d_{i,j}$). The available squared distance measurements can then be modeled in a matrix form as

$$\begin{aligned}
(\mathbf{R} \odot \mathbf{R}) \odot \mathbf{W} \;\; &= \;\; \left(\boldsymbol{\psi}\mathbf{1}_N^T - 2\mathbf{X}^T\mathbf{X} + \mathbf{1}_N\boldsymbol{\psi}^T\right) \odot \mathbf{W} \\
&= \;\; \mathbf{W}diag(\boldsymbol{\psi}) - 2\mathbf{X}^T\mathbf{X} \odot \mathbf{W} + diag(\boldsymbol{\psi})\mathbf{W} \qquad (5.4)
\end{aligned}$$

with $[\mathbf{R}]_{i,j} = [\mathbf{R}]_{j,i} = r_{i,j}$, and $\mathbf{W} \in \mathcal{R}^{N \times N}$ is the symmetric communication connectivity matrix of the network, where its (i,j)th element equal to one if the ith node and the jth node communicate with each other, and equal to zero if they do not communicate with each other. The elements on the diagonal of this matrix are arbitrary. The positions of the nodes directly appear in the second

term of (5.4). The idea is therefore to omit the first and the last terms in (5.4) while keeping the second term. We omit these two terms by pre- and post-multiplying (5.4) by an orthogonal projection matrix $\mathbf{P}_w \in \mathcal{R}^{N \times N}$ such that $\mathbf{P}_w \mathbf{W} = \mathbf{W} \mathbf{P}_w = \mathbf{0}$. This operation can be considered as an extension of the classical MDS. We assume that the structure of $\mathbf{W}$ is

$$\mathbf{W} = \left[\begin{array}{cc} \mathbf{1}_M \mathbf{1}_M^T & \mathbf{1}_M \mathbf{1}_{N-M}^T \\ \mathbf{1}_{N-M} \mathbf{1}_M^T & \mathbf{\Lambda} \end{array} \right] \tag{5.5}$$

where $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_1, \ldots, \lambda_{N-M})$. Given $\{\lambda_j\}_{j=1}^{N-M}$ we determine $\mathbf{P}_w$. Let $\tilde{\mathbf{W}}$ be the matrix containing a basis for the column span of $\mathbf{W}$. Then $\mathbf{P}_w = \mathbf{I}_N - \tilde{\mathbf{W}} \left(\tilde{\mathbf{W}}^T \tilde{\mathbf{W}} \right)^{-1} \tilde{\mathbf{W}}^T$. We discuss two possible types of $\tilde{\mathbf{W}}$, and present the effect of each projection on the node positions. We emphasize that in the proposed localization approach these projections are not directly applied to the positions. Still, the purpose is to give an intuition for choosing these projections by presenting their effect on the positions of the nodes.

5.4.1   Projection type A

Assume $\mathbf{\Lambda} = \mathbf{0}_{N-M} \mathbf{0}_{N-M}^T$, where $\mathbf{0}_n \in \mathcal{R}^{n \times 1}$ is a vector with all elements equal to zero. Then $\tilde{\mathbf{W}}^{(A)}$ and its associated orthogonal projection matrix $\mathbf{P}_w^{(A)}$ are

$$\tilde{\mathbf{W}}^{(A)} = \left[\begin{array}{cc} \mathbf{1}_M & \mathbf{1}_M \\ \mathbf{1}_{N-M} & \mathbf{0}_{N-M} \end{array} \right] \in \mathcal{R}^{N \times 2} \tag{5.6}$$

$$\mathbf{P}_w^{(A)} = \left[\begin{array}{cc} \mathbf{I}_M - \frac{1}{M} \mathbf{1}_M \mathbf{1}_M^T & \mathbf{0}_M \mathbf{0}_{N-M}^T \\ \mathbf{0}_{N-M} \mathbf{0}_M^T & \mathbf{I}_{N-M} - \frac{1}{N-M} \mathbf{1}_{N-M} \mathbf{1}_{N-M}^T \end{array} \right] \tag{5.7}$$

The projected node positions are $\mathbf{P}_w^{(A)} \mathbf{X}^T = \left[\begin{array}{cc} \tilde{\mathbf{X}}_1^T & \tilde{\mathbf{X}}_2^T \end{array} \right]^T$ where

$$\tilde{\mathbf{X}}_1 = \mathbf{X}_1 - \mathbf{x}_1^{(c)} \mathbf{1}_M^T \tag{5.8}$$

$$\tilde{\mathbf{X}}_2 = \mathbf{X}_2 - \mathbf{x}_2^{(c)} \mathbf{1}_{N-M}^T \tag{5.9}$$

with $\mathbf{x}_1^{(c)} = \frac{1}{M} \sum_{j=1}^M \mathbf{x}_j$ and $\mathbf{x}_2^{(c)} = \frac{1}{N-M} \sum_{j=M+1}^N \mathbf{x}_j$ the centers of gravity of the first and second node group, respectively. The advantage of this projection is that it decouples the two groups, but the disadvantage is that both groups are translated to the origin. To obtain the relative distance between the centers of gravity of the two groups we need the following projection.

5.4.2 Projection type B

Assume $\mathbf{\Lambda} = (N-M)\mathbf{I}_{N-M}$. Then $\tilde{\mathbf{W}}^{(B)}$ and its associated orthogonal projection matrix $\mathbf{P}_w^{(B)}$ are

$$\tilde{\mathbf{W}}^{(B)} = \left[\begin{array}{c} \mathbf{1}_M \mathbf{1}_{N-M}^T \\ \mathbf{\Lambda} \end{array} \right] \in \mathcal{R}^{N \times (N-M)} \tag{5.10}$$

$$\mathbf{P}_w^{(A)} = \left[\begin{array}{cc} \mathbf{I}_M - \frac{1}{N}\mathbf{1}_M\mathbf{1}_M^T & -\frac{1}{N}\mathbf{1}_M\mathbf{1}_{N-M}^T \\ -\frac{1}{N}\mathbf{1}_{N-M}\mathbf{1}_M^T & \frac{M}{(N-M)N}\mathbf{1}_{N-M}\mathbf{1}_{N-M}^T \end{array} \right] \tag{5.11}$$

The projected node positions are

$$\mathbf{P}_w^{(B)}\mathbf{X}^T = \left[\begin{array}{cc} \mathbf{I}_M & \mathbf{q}_1 \\ \mathbf{0}_{N-M}\mathbf{0}_M^T & -\mathbf{q}_2 \end{array} \right] \left[\begin{array}{c} \tilde{\mathbf{X}}_1^T \\ \mathbf{u}^T \end{array} \right] \tag{5.12}$$

where the relative distance vector between the centers of gravity of the two groups is defined as

$$\mathbf{u} = \mathbf{x}_1^{(c)} - \mathbf{x}_2^{(c)} \tag{5.13}$$

and $\mathbf{q}_1 = \frac{N-M}{N}\mathbf{1}_M$, $\mathbf{q}_2 = \frac{M}{N}\mathbf{1}_{N-M}$. The advantage of using this projection is that it retains the relative translation between the centers of gravity of the two groups, while the disadvantage is that all the nodes of the second group are translated to the same position.

5.5 The proposed localization approach

The idea of reconstructing the configuration of the sensor network is as follows: use $\mathbf{P}_w^{(A)}$ to estimate $\tilde{\mathbf{X}}_1$, and $\tilde{\mathbf{X}}_2$. If one is also interested in the relative translation vector $\mathbf{u}$, then first use $\mathbf{P}_w^{(B)}$ to estimate $\mathbf{u}$ (given the estimate of $\tilde{\mathbf{X}}_1$), and then update the estimate of $\tilde{\mathbf{X}}_2$ by $\tilde{\mathbf{X}}_2' = \tilde{\mathbf{X}}_2 + \mathbf{u}$. We now discuss each of these steps in detail.

5.5.1 The result of using projection $\mathbf{P}_w^{(A)}$

By post- and pre- multiplying the left hand side of (5.4) by $\frac{1}{\sqrt{2}}\mathbf{P}_w^{(A)}$ we achieve a measurement matrix $\mathbf{B}$, which is defined as

$$\mathbf{B} \triangleq \left[\begin{array}{cc} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^T & \mathbf{B}_{22} \end{array} \right] = -\frac{1}{2}\mathbf{P}_w^{(A)} \left((\mathbf{R} \odot \mathbf{R}) \odot \mathbf{W} \right) \mathbf{P}_w^{(A)}. \tag{5.14}$$

Meanwhile, by post- and pre- multiplying the right hand side of (5.4) by $\frac{1}{\sqrt{2}}\mathbf{P}_w^{(A)}$, and making use of $\mathbf{P}_w^{(A)}\mathbf{W} = \mathbf{W}\mathbf{P}_w^{(A)} = \mathbf{0}$, we arrive at

$$\mathbf{P}_w^{(A)}(\mathbf{X}^T\mathbf{X} \odot \mathbf{W})\mathbf{P}_w^{(A)} = \begin{bmatrix} \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 & \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_2 \\ \tilde{\mathbf{X}}_2^T\tilde{\mathbf{X}}_1 & \mathbf{0}_{N-M}\mathbf{0}_{N-M}^T \end{bmatrix} \tag{5.15}$$

As a result, in the noiseless case, we can model $\mathbf{B}$ as

$$\begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{12}^T & \mathbf{B}_{22} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 & \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_2 \\ \tilde{\mathbf{X}}_2^T\tilde{\mathbf{X}}_1 & \mathbf{0}_{N-M}\mathbf{0}_{N-M}^T \end{bmatrix} \tag{5.16}$$

Due to the noise, we recover $\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2$ (up to rotation, reflection and translation) as $[\hat{\tilde{\mathbf{X}}}_1, \hat{\tilde{\mathbf{X}}}_2] = \mathrm{argmin}_{[\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2]} J(\tilde{\mathbf{X}}_1, \tilde{\mathbf{X}}_2)$ where

$$\begin{aligned} J(\mathbf{X}) &= \left\| \mathbf{B} - \mathbf{P}_w^{(A)}(\mathbf{X}^T\mathbf{X} \odot \mathbf{W})\mathbf{P}_w^{(A)} \right\|_F^2 \\ &= \left\| \mathbf{B}_{11} - \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 \right\|_F^2 + 2\left\| \mathbf{B}_{12} - \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_2 \right\|_F^2 \end{aligned} \tag{5.17}$$

It is clear that $\tilde{\mathbf{X}}_2$ which minimizes (5.17) is

$$\hat{\tilde{\mathbf{X}}}_2 = \left(\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T \right)^{-1} \tilde{\mathbf{X}}_1\mathbf{B}_{12} \tag{5.18}$$

Substituting (5.18) into (5.17) yields

$$\begin{aligned} J(\tilde{\mathbf{X}}_1) &= \left\| \mathbf{B}_{11} - \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 \right\|_F^2 + 2\left\| \mathbf{B}_{12} - \tilde{\mathbf{X}}_1^T\left(\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T \right)^{-1}\tilde{\mathbf{X}}_1\mathbf{B}_{12} \right\|_F^2 \quad (5.19) \\ &= \mathrm{tr}\left(\mathbf{B}_{11}^T\mathbf{B}_{11} + 2\mathbf{B}_{12}^T\mathbf{B}_{12} - 2\mathbf{B}_{11}\tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 \right. \\ &\qquad \left. + \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 - 2\mathbf{B}_{12}^T\tilde{\mathbf{X}}_1^T\left(\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T \right)^{-1}\tilde{\mathbf{X}}_1\mathbf{B}_{12} \right) \end{aligned}$$

By taking the derivative with respect to (w.r.t.) $\tilde{\mathbf{X}}_1^T$ we obtain that

$$\begin{aligned} \frac{\partial J(\tilde{\mathbf{X}}_1)}{\partial \tilde{\mathbf{X}}_1^T} &= -\mathbf{B}_{11}\tilde{\mathbf{X}}_1^T + \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T + \tilde{\mathbf{X}}_1^T\left(\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T \right)^{-1}\tilde{\mathbf{X}}_1\mathbf{B}_{12}\mathbf{B}_{12}^T\tilde{\mathbf{X}}_1^T\left(\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T \right)^{-1} \\ &\quad - \mathbf{B}_{12}\mathbf{B}_{12}^T\tilde{\mathbf{X}}_1^T\left(\tilde{\mathbf{X}}_1\tilde{\mathbf{X}}_1^T \right)^{-1} \end{aligned} \tag{5.20}$$

Since no closed-form expression for $\hat{\tilde{\mathbf{X}}}_1$ that zeros (5.20) exists, we estimate it as detailed in Algorithm . The final estimate is denoted by $\hat{\tilde{\mathbf{X}}}_1$. After obtaining $\hat{\tilde{\mathbf{X}}}_1$, we substitute the result in (5.18) and obtain $\hat{\tilde{\mathbf{X}}}_2$.

Algorithm Estimating the positions of the fully connected nodes

- Initial step1: Construct $\mathbf{B}$ according to (5.14).

- Initial step2: $\hat{\tilde{\mathbf{X}}}_1^{(MDS)} = \mathbf{\Phi}^{1/2}\mathbf{C}^T$, where $\mathbf{\Phi} \in \mathcal{R}^{2\times 2}$, $\mathbf{C} \in \mathcal{R}^{M\times 2}$ contain the two largest eigenvalues of $\mathbf{B}_{11}$, and their two associated orthonormal eigenvectors, respectively.

- Updating step: Given $\hat{\tilde{\mathbf{X}}}_1^{(MDS)}$ determine $\hat{\tilde{\mathbf{X}}}_1$ as follows:

  1. Let $\hat{\tilde{\mathbf{X}}}_1^{(k)}$ be the value of $\hat{\tilde{\mathbf{X}}}_1$ at the kth iteration step.

  2. Calculate the gradient $\mathbf{G}^{(k)} = \left.\frac{\partial J(\tilde{\mathbf{X}}_1)}{\partial \tilde{\mathbf{X}}_1^T}\right|_{\hat{\tilde{\mathbf{X}}}_1 = \hat{\tilde{\mathbf{X}}}_1^{(k)}}$ according to (5.20).

  3. Update the step size, $\mu^{(k)}$, as $\mu^{(k)} = \text{argmin}_\mu J\left(\hat{\tilde{\mathbf{X}}}_1^{(k)} - \mu\mathbf{G}^{(k)}\right)$ using (5.19).

  4. Update the estimate, $\hat{\tilde{\mathbf{X}}}_1^{(k+1)} = \hat{\tilde{\mathbf{X}}}_1^{(k)} - \mu^{(k)}\mathbf{G}^{(k)}$.

  5. if $\left|J\left(\hat{\tilde{\mathbf{X}}}_1^{(k+1)}\right) - J\left(\hat{\tilde{\mathbf{X}}}_1^{(k)}\right)\right| < \epsilon$ where ϵ is a predefined tolerance, then $\hat{\tilde{\mathbf{X}}}_1 = \hat{\tilde{\mathbf{X}}}_1^{(k+1)}$, else perform steps 1-4.

5.5.2 The result of using projection $\mathbf{P}_w^{(B)}$

Given the estimated positions of the previous step, we are interested in estimating the relative distance vector between the two groups in order to reconstruct the complete network configuration. Similarly as in the previous subsection, by post- and pre- multiplying the left hand side of (5.4) by $\frac{1}{\sqrt{2}}\mathbf{P}_w^{(B)}$, we obtain a measurement matrix $\mathbf{F}$, which is defined as

$$\mathbf{F} = \left[\begin{array}{cc} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{12}^T & \mathbf{F}_{22} \end{array}\right] = -\frac{1}{2}\mathbf{P}_w^{(B)}\left((\mathbf{R} \odot \mathbf{R}) \odot \mathbf{W}\right)\mathbf{P}_w^{(B)} \qquad (5.21)$$

At the same time, by post- and pre- multiplying the right hand side of (5.4) by $\frac{1}{\sqrt{2}}\mathbf{P}_w^{(B)}$, we arrive at

$$\mathbf{P}_w^{(B)}(\mathbf{X}^T\mathbf{X} \odot \mathbf{W})\mathbf{P}_w^{(B)} = \left[\begin{array}{cc} \tilde{\mathbf{X}}_1^T\tilde{\mathbf{X}}_1 + \mathbf{q}_1\mathbf{u}^T\tilde{\mathbf{X}}_1 + \tilde{\mathbf{X}}_1^T\mathbf{u}\mathbf{q}_1^T & -\tilde{\mathbf{X}}_1^T\mathbf{u}\mathbf{q}_2^T \\ -\mathbf{q}_2\mathbf{u}^T\tilde{\mathbf{X}}_1 & \mathbf{0}_{N-M}\mathbf{0}_{N-M}^T \end{array}\right]$$
$$+\gamma\mathbf{q}\mathbf{q}^T \qquad (5.22)$$

where $\mathbf{q} = \left[\begin{array}{cc} \mathbf{q}_1^T & -\mathbf{q}_2^T \end{array}\right]^T$, $\gamma = \|\mathbf{u}\|^2 + \eta^2$, and $\eta^2 = \frac{1}{N-M}\sum_{j=m+1}^N \|\mathbf{x}_j\|^2 - \left\|\mathbf{x}_2^{(c)}\right\|^2$. Note that η^2 describes the scattering radius around the center of gravity

of the positions of the nodes in the second group. Then in the noiseless case, $\mathbf{F}$ can be modeled as

$$
\begin{bmatrix} \mathbf{F}_{11} & \mathbf{F}_{12} \\ \mathbf{F}_{12}^T & \mathbf{F}_{22} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{X}}_1^T \tilde{\mathbf{X}}_1 + \mathbf{q}_1 \mathbf{u}^T \tilde{\mathbf{X}}_1 + \tilde{\mathbf{X}}_1^T \mathbf{u} \mathbf{q}_1^T & -\tilde{\mathbf{X}}_1^T \mathbf{u} \mathbf{q}_2^T \\ -\mathbf{q}_2 \mathbf{u}^T \tilde{\mathbf{X}}_1 & \mathbf{0}_{N-M} \mathbf{0}_{N-M}^T \end{bmatrix} + \gamma \mathbf{q} \mathbf{q}^T
$$
(5.23)

We are interested in estimating the relative distance $\mathbf{u}$ given the translated positions of the first group $\tilde{\mathbf{X}}_1$, and consider γ as a nuisance parameter. Due to the noise, we estimate $\mathbf{u}$ by minimizing

$$
H(\mathbf{u}, \gamma) = \left\| \mathbf{F}_{11} - \left(\tilde{\mathbf{X}}_1^T \tilde{\mathbf{X}}_1 + \mathbf{q}_1 \mathbf{u}^T \tilde{\mathbf{X}}_1 + \tilde{\mathbf{X}}_1^T \mathbf{u} \mathbf{q}_1^T + \gamma \mathbf{q}_1 \mathbf{q}_1^T \right) \right\|_F^2
$$
$$
+ 2 \left\| \mathbf{F}_{12} - \left(\tilde{\mathbf{X}}_1^T \mathbf{u} \mathbf{q}_2^T - \gamma \mathbf{q}_1 \mathbf{q}_2^T \right) \right\|_F^2 + \left\| \mathbf{F}_{22} - \gamma \mathbf{q}_2 \mathbf{q}_2^T \right\|_F^2 \quad (5.24)
$$

By taking the derivative w.r.t. $\mathbf{u}$ and equating the result to zero, and using the fact that $\frac{N-M}{N} \mathbf{1}_M^T \tilde{\mathbf{X}}_1^T \mathbf{u} = 0$ (since the center of gravity of $\tilde{\mathbf{X}}_1$ is in the origin), we get that the estimated relative distance vector is

$$
\hat{\mathbf{u}} = \left(\tilde{\mathbf{X}}_1 \tilde{\mathbf{X}}_1^T \right)^{-1} \tilde{\mathbf{X}}_1 \left(\frac{1}{M} \mathbf{F}_{11} \mathbf{1}_M - \frac{1}{N-M} \mathbf{F}_{12} \mathbf{1}_{N-M} \right)
$$
(5.25)

Note that $\hat{\mathbf{u}}$ does not depend on the estimate of the nuisance parameter γ, and we therefore do not proceed in estimating it. We now substitute $\hat{\tilde{\mathbf{X}}}_1$ (obtained in the first step) instead of $\tilde{\mathbf{X}}_1$ in (5.25). We then update $\hat{\tilde{\mathbf{X}}}$ (obtained in the first step) by $\hat{\tilde{\mathbf{X}}}_2' = \hat{\tilde{\mathbf{X}}}_2 + \hat{\mathbf{u}} \mathbf{1}_{N-M}^T$. Finally, $\hat{\mathbf{X}} = \begin{bmatrix} \hat{\tilde{\mathbf{X}}}_1 & \hat{\tilde{\mathbf{X}}}_2' \end{bmatrix} \in \mathcal{R}^{2 \times N}$ contains the estimated node positions up to rotation, reflection, and translation w.r.t. the original configuration (which can be corrected using anchors in the network).

5.6   Simulation results

We compare our proposed algorithm with the MDS-MAP method [110], the SVD-Reconstruct method [113] (where each missing entry is replaced by zero since it is the optimal choice as indicated in [113]), and with Nyström's method [114]. We also show some results for the SMACOF algorithm [112, pp. 150-157], one for a random initial point and one for an improved initial point (our proposed method). We consider a square area of 100×100 $[m \times m]$, and $N = 100$ nodes randomly positioned. The tolerance of the iterative search in the first estimation step is $\epsilon = 0.005$.

In the first simulation, we consider the cases of $M = 10$, $M = 50$ and $M = 90$. For each case we varied the noise variance from $\sigma^2 = 0.2$ to $\sigma^2 = 5$. For

each noise variance we consider $K = 50$ configuration realizations, and for each configuration we perform $N_{exp} = 50$ Monte-Carlo (MC) trials. Since all positions are unknown, we define the RMSE as $\bar{\varepsilon} = \sqrt{\frac{1}{KN_{exp}} \sum_{k=1}^{KN_{exp}} \bar{\varepsilon}_k^2}[m]$, where $\bar{\varepsilon}_k = \sqrt{\frac{1}{N_d} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} w_{i,j} \left(\hat{d}_{i,j} - d_{i,j} \right)^2}[m]$, with $\hat{d}_{i,j}$ the estimated distance between the connected nodes i and j at one configuration and one MC trial, and N_d is the number of connections (in this case $N_d = \frac{1}{2}M(N-1)$). The results are shown in Fig. 5.1. As can be seen, our proposed algorithm has the smallest RMSE, compared to other MDS-like solutions. Both MDS-MAP and SVD-Reconstruct have worse performance when the number of connections is small ($M = 10$), however, when the number of the connections is large ($M = 90$), MDS-MAP has a similar performance as the proposed algorithm and the Nyström's method. From the results, we observe that a good initial point for the SMACOF algorithm is crucial and it is clear that the better the initial point, the faster the convergence of SMACOF. That is why our method is preferred over other MDS-like solutions if an initial point for SMACOF is sought for. In Fig. 5.2 (left plot) we compare the RMSE of $\hat{\bar{\mathbf{X}}}_1$ with that of the initial estimate $\hat{\bar{\mathbf{X}}}_1^{(MDS)}$ (in this case $N_d = M(M-1)/2$), and in Fig. 5.2 (right plot) we compare the RMSE of $\hat{\bar{\mathbf{X}}}_2'$ of the proposed method with that of Nyström's method (in this case $N_d = \frac{1}{2}M(N-M)$). For both cases we assume that $M = 10$. As can be seen, when the noise variance increases, the gap between the RMSEs of $\hat{\bar{\mathbf{X}}}_1$ and $\hat{\bar{\mathbf{X}}}_1^{(MDS)}$ increases, and thus the iterative solution improves the RMSE of the initial estimate. The RMSE of $\hat{\bar{\mathbf{X}}}_2'$ is also improved w.r.t. the Nyström's method. So, our algorithm outperforms this algorithm for all values of σ^2.

In the second simulation, we compare the RMSE versus M/N of our proposed algorithm with the other methods. We varied M from 10 to 90 with a step of 5. We assume that $\sigma^2 = 0.2[m^2]$. The results are shown in Fig. 5.3. As can be seen, our algorithm outperforms the other MDS-like methods for all values of M/N.

5.7   Conclusions

We consider the problem of reconstructing the configuration of a sensor network (up to rotation, reflection, and translation) from pairwise distance measurements assuming the network is composed of two groups: one group contains nodes that communicate with each other, and the second group contains nodes that do not communicate with each other, and only communicate with each of the nodes in the first group. The classical MDS algorithm cannot be applied in this case. Our approach is based on performing two projections on the available set of distance measurements. Simulations show that the proposed algorithm outperforms other methods based on matrix completion techniques. Future work will focus on: 1) examining other projections and their effect on the node positions; 2) analyzing
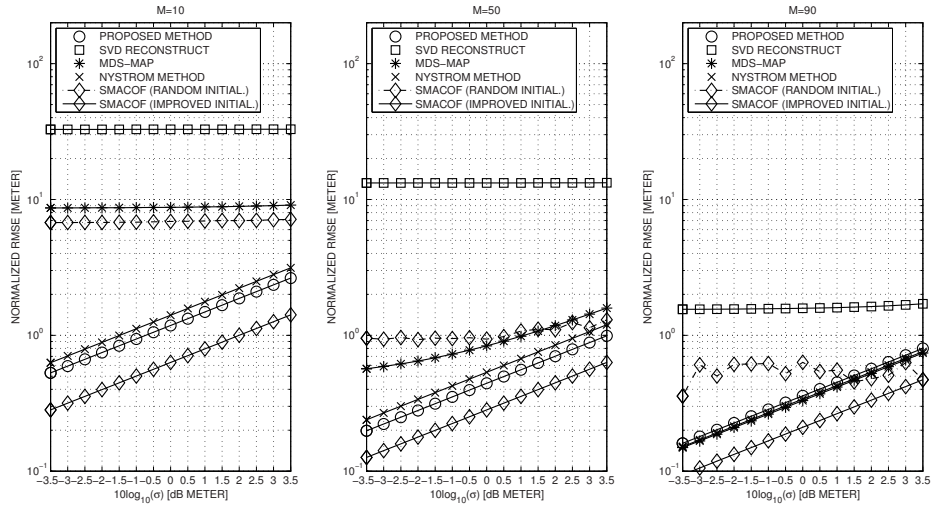
Figure 5.1: RMSEs versus the noise variance for $M = 10$ (left), $M = 50$ (center), and $M = 90$ (right).

the performance of the first step; 3) evaluating the complexity of the proposed approach.
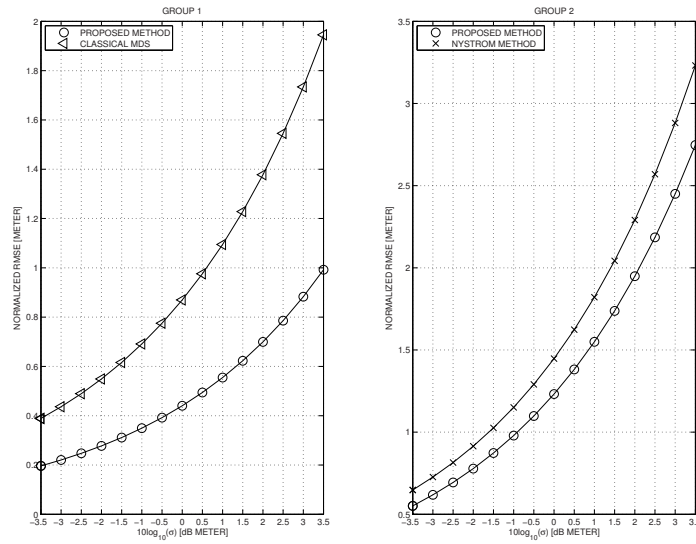
Figure 5.2: RMSEs of the initial solution and the iterative solution of the first group (left), and the RMSE of the second group using the proposed method and Nyström's method (right).
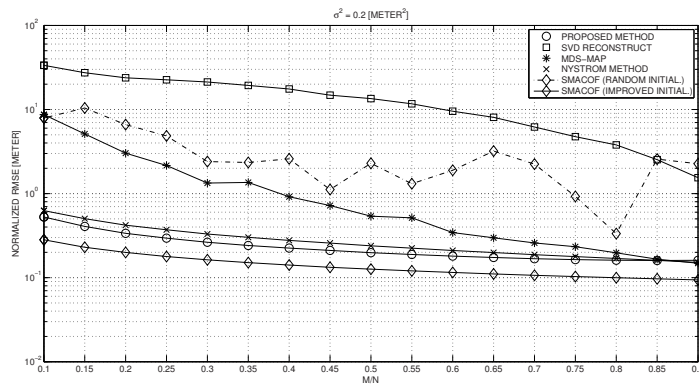
Figure 5.3: RMSEs versus the ratio M/N for $\sigma^2 = 0.2$.

CHAPTER **6**

Reference-free Time-based Localization

6.1   Introduction

Closed-form localization solutions based on TOAs or TDOAs are used to locate a target node with the help of anchors (nodes with known positions). They are appreciated for real-time localization applications, initiating iterative localization algorithms, and facilitating Kalman tracking [116]. They have much lower complexity compared to the optimal maximum likelihood estimator (MLE), and also do not require prior knowledge of noise statistics. However, a common feature of existing closed-form localization solutions is reference dependency. The reference here indicates the time associated with the reference anchor. For instance, in order to measure TDOAs, a reference anchor has to be chosen first [9]. The reference anchor is also needed to cancel nuisance parameters in closed-form solutions based on TOAs or TDOAs [117]. Thus, the localization performance depends heavily on the reference selection. There are some efforts to improve the reference selection [87, 118, 119], but they mainly rely on heuristics. Furthermore, when TOAs are measured using the one-way ranging protocol for calculating the distance between the target and the anchor, stringent synchronization is required between these two nodes in the conventional methods [9, 11]. However, it is difficult to maintain synchronization due to the clock inaccuracy and other error sources. Therefore, various closed-form localization methods resort to using TDOA measurements to relax this synchronization constraint between the target and the anchor. These methods only require synchronization among the anchors, e.g., the source localization methods based on TDOAs using a passive

_____

The results in this chapter appeared in [115]

sensor array* [6, 120–123].

In this chapter, we also relax the above synchronization requirement, and consider a scenario, where anchor nodes are synchronized, and the clock of the target node runs freely. However, instead of using TDOAs, we model the asynchronous effect as a common bias, and propose reference-free least-squares (LS), weighted LS (WLS), and constrained WLS (CWLS) localization estimators based on TOA measurements. Furthermore, we generalize existing reference-based localization solutions using TOA or TDOA measurements, which are scattered over different research areas, and provide new insights into their relations, which have been overlooked. We clarify that the reference dependency for reference-based WLS estimators using TOA measurements can be decoupled by the optimal weighting matrix, which also makes all those estimators identical. However, the influence of the reference selection for reference-based WLS estimators using a nonredundant set of TDOA measurements cannot be compensated by the optimal weighting matrix. But the optimal weighting matrix can make the estimators using the same set equivalent as well. Moreover, the Cramér-Rao bounds (CRBs) are derived as benchmarks for comparison.

The rest of this chapter is organized as follows. In Section 6.2, different kinds of reference-free TOA-based estimators are proposed, as well as existing reference-based estimators using TOA measurements. Their relations are thoroughly investigated. In Section 6.3, we generalize existing reference-based localization algorithms using TDOA measurements, and shed light on their relations as well. Simulation results and performance bounds are shown in Section 6.4. Conclusions are drawn at the end of the chapter.

6.2   Localization based on TOA measurements

Considering M anchor nodes and one target node, we would like to estimate the position of the target node. We recall $\mathbf{X}_a$ and $\mathbf{x}$ defined in Chapter III, which indicate the known coordinates of the anchor nodes and the unknown coordinates of the target node, respectively. Our method can also be extended for multiple target nodes. We remark that in a large scale WSN, it is common to localize target nodes in a sequential way [27]. Therefore, the multiple-anchors-one-target scenario here is of practical interest. We can even consider a case with a moving anchor, in which a ranging signal is periodically transmitted by the target node, and all the positions where the moving anchor receives the ranging signal are viewed as the fixed positions of some virtual anchors. We assume that all the anchors are synchronized, and their clock skews are equal to 1, whereas the clock of the target node runs freely. Furthermore, we assume that the target node transmits a ranging signal, and all the anchors act as receivers. We remark that other systems may share the same data model such as a passive sensor array for source localization or a GPS system, where a GPS receiver locates itself by exploring the received

---

*The sensor elements of a passive sensor array are equivalent to the anchor nodes here.

ranging signals from several satellites [3]. All the satellites are synchronized to an atomic clock, but the GPS receiver has a clock offset relative to the satellite clock. Note that this is a stricter synchronization requirement than ours, as we allow the clock of the target node to run freely. Every satellite sends a ranging signal and a corresponding transmission time. The GPS receiver measures the TOAs, and calculates the time-of-flight (TOF) plus an unknown offset. In this section, TOA measurements are used, and TDOA measurements are employed in the next section.

6.2.1   System model

In this section, all localization algorithms are based on TOA measurements. When the target node transmits a ranging signal, all the anchors receive it and record a timestamp upon the arrival of the ranging signal independently. We define a vector $\mathbf{u}$ of length M to collect all the distances corresponding to the timestamps, which is given by $\mathbf{u} = [u_1,\ u_2,\ \ldots,\ u_M]^T$. We employ b to denote the distance corresponding to the true target node transmission instant, which is unknown. We remark that if we consider a GPS system, then $\mathbf{u}$ collects the distances corresponding to the biased TOFs calculated by the GPS receiver, and b indicates the distance bias corresponding to the unknown clock offset of the GPS receiver relative to the satellite. Consequently, the TOA measurements can be modeled as

$$\mathbf{u} - b\mathbf{1}_M \quad = \quad \mathbf{d} + \mathbf{n}, \tag{6.1}$$

where $\mathbf{d} = [d_1,\ d_2,\ \ldots,\ d_M]^T$, with $d_i = \|\mathbf{x}_i - \mathbf{x}\|$ the true distance between the ith anchor node and the target node, and $\mathbf{n} = [n_1,\ n_2,\ \ldots,\ n_M]^T$ with n_i the distance error term corresponding to the TOA measurement error at the ith anchor, which can be modeled as a random variable with zero mean and variance σ_i^2, and which is independent of the other terms ($E[n_i n_j] = 0, i \neq j$). We remark that instead of using TDOAs to directly get rid of the distance bias, we use TOAs and take the bias into account in the system model.

6.2.2   Localization Based on Squared TOA Measurements

Proposed localization algorithms

Note that (6.1) is a nonlinear equation with respect to (w.r.t.) $\mathbf{x}$. To solve it, a maximum likelihood estimator (MLE) can be derived, which is optimal in the sense that for a large number of data it is unbiased and approaches the CRB. However, the MLE has a high computational complexity, and also requires the unknown noise statistics. Therefore, low-complexity solutions are of great interest for localization. From $\|\mathbf{x}_i - \mathbf{x}\|^2 = \|\mathbf{x}_i\|^2 - 2\mathbf{x}_i^T\mathbf{x} + \|\mathbf{x}\|^2$, we derive that $\mathbf{d} \odot \mathbf{d} = \boldsymbol{\psi}_a - 2\mathbf{X}_a^T\mathbf{x} + \|\mathbf{x}\|^2\mathbf{1}_M$, where $\boldsymbol{\psi}_a = [\|\mathbf{x}_1\|^2,\ \|\mathbf{x}_2\|^2,\ \ldots,\ \|\mathbf{x}_M\|^2]^T$.

Element-wise multiplication at both sides of (6.1) is carried out, which leads to

$$\mathbf{u} \odot \mathbf{u} - 2b\mathbf{u} + b^2 \mathbf{1}_M \quad = \quad \boldsymbol{\psi}_a - 2\mathbf{X}_a^T \mathbf{x} + \|\mathbf{x}\|^2 \mathbf{1}_M + 2\mathbf{d} \odot \mathbf{n} + \mathbf{n} \odot \mathbf{n}. \quad (6.2)$$

Moving knowns to one side and unknowns to the other side, we achieve

$$\boldsymbol{\psi}_a - \mathbf{u} \odot \mathbf{u} \quad = \quad 2\mathbf{X}_a^T \mathbf{x} - 2b\mathbf{u} + (b^2 - \|\mathbf{x}\|^2)\mathbf{1}_M + \mathbf{m}, \quad (6.3)$$

where $\mathbf{m} = -(2\mathbf{d} \odot \mathbf{n} + \mathbf{n} \odot \mathbf{n})$. The stochastic properties of $\mathbf{m}$ are as follows

$$
\begin{aligned}
E[[\mathbf{m}]_i] &= -\sigma_i^2 \approx 0, & (6.4) \\
[\boldsymbol{\Sigma}]_{i,j} &= E[[\mathbf{m}]_i[\mathbf{m}]_j] - E[[\mathbf{m}]_i]E[[\mathbf{m}]_j] \\
&= E[(2d_i n_i + n_i^2)(2d_j n_j + n_j^2)] - \sigma_i^2 \sigma_j^2 \\
&= 4d_i d_j E[n_i n_j] + E[n_i^2 n_j^2] - \sigma_i^2 \sigma_j^2 \\
&= \begin{cases} 4d_i^2 \sigma_i^2 + 2\sigma_i^4 \approx 4d_i^2 \sigma_i^2, & i = j \\ 0, & i \neq j \end{cases}, & (6.5)
\end{aligned}
$$

where we ignore the higher order noise terms to obtain (6.5) and assume that the noise mean $E[[\mathbf{m}]_i] \approx 0$ under the condition of sufficiently small measurement errors. Note that the noise covariance matrix $\boldsymbol{\Sigma}$ depends on the unknown $\mathbf{d}$.

Defining $\boldsymbol{\phi} = \boldsymbol{\psi}_a - \mathbf{u} \odot \mathbf{u}$, $\mathbf{y} = [\mathbf{x}^T, \ b, \ b^2 - \|\mathbf{x}\|^2]^T$, and $\mathbf{A} = [2\mathbf{X}_a^T, \ -2\mathbf{u}, \ \mathbf{1}_M]$, we can finally rewrite (6.3) as

$$\boldsymbol{\phi} \quad = \quad \mathbf{A}\mathbf{y} + \mathbf{m}. \quad (6.6)$$

Ignoring the parameter relations in $\mathbf{y}$, an unconstrained LS and WLS estimate of $\mathbf{y}$ can be computed respectively given by

$$\hat{\mathbf{y}} \quad = \quad (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \boldsymbol{\phi}, \quad (6.7)$$

and

$$\hat{\mathbf{y}} \quad = \quad (\mathbf{A}^T \mathbf{W} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{W} \boldsymbol{\phi}, \quad (6.8)$$

where $\mathbf{W}$ is a weighting matrix of size $M \times M$. Note that $M \geq l + 2$ is required in (6.7) and (6.8), which indicates that we need at least four anchors to estimate the target position on a plane. The optimal $\mathbf{W}$ is $\mathbf{W}^* = \boldsymbol{\Sigma}^{-1}$, which depends on the unknown $\mathbf{d}$ as we mentioned before. Thus, we can update it iteratively, and the resulting iterative WLS can be summarized as follows:

1. Initialize $\mathbf{W}$ using the estimate of $\mathbf{d}$ based on the LS estimate of $\mathbf{x}$;

2. Estimate $\hat{\mathbf{y}}$ using (6.8);

3. Update $\mathbf{W} = \hat{\boldsymbol{\Sigma}}^{-1}$, where $\hat{\boldsymbol{\Sigma}}$ is computed using $\hat{\mathbf{y}}$;

4. Repeat Steps 2) and 3) until a stopping criterion is satisfied.

The typical stopping criteria are discussed in [124]. We stop the iterations when $\|\hat{\mathbf{y}}^{(k+1)} - \hat{\mathbf{y}}^{(k)}\| \leq \epsilon$, where $\hat{\mathbf{y}}^{(k)}$ is the estimate of the kth iteration and ϵ is a given threshold [125]. An estimate of $\mathbf{x}$ is finally given by

$$\hat{\mathbf{x}} = [\mathbf{I}_l \quad \mathbf{0}_{l \times 2}]\hat{\mathbf{y}}. \tag{6.9}$$

To accurately estimate $\mathbf{y}$, we can further explore the relations among the parameters in $\mathbf{y}$. A constrained weighted LS (CWLS) estimator is obtained as

$$\hat{\mathbf{y}} = \arg\min_{\hat{\mathbf{y}}} (\boldsymbol{\phi} - \mathbf{A}\mathbf{y})^T \mathbf{W} (\boldsymbol{\phi} - \mathbf{A}\mathbf{y}) \tag{6.10}$$

subject to

$$\mathbf{y}^T \mathbf{J} \mathbf{y} + \boldsymbol{\rho}^T \mathbf{y} = 0, \tag{6.11}$$

where $\boldsymbol{\rho} = [\mathbf{0}_{l+1}^T \quad 1]^T$ and

$$\mathbf{J} = \begin{bmatrix} \mathbf{I}_l & \mathbf{0}_l & 0 \\ \mathbf{0}_l^T & -1 & 0 \\ \mathbf{0}_l^T & 0 & 0 \end{bmatrix}. \tag{6.12}$$

Solving the CWLS problem is equivalent to minimizing the Lagrangian [6, 11]

$$\mathcal{L}(\mathbf{y}, \lambda) = (\boldsymbol{\phi} - \mathbf{A}\mathbf{y})^T \mathbf{W} (\boldsymbol{\phi} - \mathbf{A}\mathbf{y}) + \lambda(\mathbf{y}^T \mathbf{J} \mathbf{y} + \boldsymbol{\rho}^T \mathbf{y}), \tag{6.13}$$

where λ is a Lagrangian multiplier. A minimum point for (6.13) is given by

$$\hat{\mathbf{y}} = (\mathbf{A}^T \mathbf{W} \mathbf{A} + \lambda \mathbf{J})^{-1} (\mathbf{A}^T \mathbf{W} \boldsymbol{\phi} - \frac{\lambda}{2} \boldsymbol{\rho}), \tag{6.14}$$

where λ is determined by plugging (6.14) into the following equation

$$\hat{\mathbf{y}}^T \mathbf{J} \hat{\mathbf{y}} + \boldsymbol{\rho}^T \hat{\mathbf{y}} = 0. \tag{6.15}$$

We could find all the seven roots of (6.15) as in [6, 11], or employ a bisection algorithm as in [126] to look for λ instead of finding all the roots. If we obtain seven roots as in [6, 11], we discard the complex roots, and plug the real roots into (6.14). Finally, we choose the estimate $\hat{\mathbf{y}}$, which fulfills (6.10). The details of solving (6.15) are mentioned in Appendix 6.A. Note that the proposed CWLS estimator (6.14) is different from the estimators in [6,11]. The CLS estimator in [6] is based on TDOA measurements, and the CWLS estimator in [11] is based on TOA measurements for a synchronous target ($b = 0$). Furthermore, we remark that the WLS estimator proposed in [127] based on the same data model as (6.1), is labeled as an extension of Bancroft's algorithm [128], which is actually

similar to the spherical-intersection (SX) method proposed in [129] for TDOA measurements. It first solves a quadratic equation in $b^2 - \|\mathbf{x}\|^2$, and then estimates $\mathbf{x}$ and b via a WLS estimator. However, it fails to provide a solution for the quadratic equation under certain circumstances, and performs unsatisfactorily when the target node is far away from the anchors [129].

Many research works have focused on LS solutions ignoring the constraint (6.11) in order to obtain low-complexity closed-form estimates [9]. As squared range (SR) measurements are employed, we call them unconstrained squared-range-based LS (USR-LS) approaches, to be consistent with [126]. Because only $\mathbf{x}$ is of interest, b and $b^2 - \|\mathbf{x}\|^2$ are nuisance parameters. Different methods have been proposed to get rid of them instead of estimating them. A common characteristic of all these methods is that they have to choose a reference anchor first, and thus we label them reference-based USR-LS (REFB-USR-LS) approaches. As a result, the performance of these REFB-USR-LS methods depends on the reference selection [9]. However, note that the unconstrained LS estimate of $\mathbf{y}$ in (6.7) does not depend on the reference selection. Thus, we call (6.7) the reference-free USR-LS (REFF-USR-LS) estimate, (6.8) the REFF-USR-WLS, and (6.14) the REFF-SR-CWLS estimate.

Moreover, we propose the subspace minimization (SM) method [123] to achieve a REFF-USR-LS estimate of $\mathbf{x}$ alone, which is identical to $\hat{\mathbf{x}}$ in (6.7), but shows more insight into the links among different estimators. Treating b and $b^2 - \|\mathbf{x}\|^2$ as nuisance parameters, we try to get rid of them by orthogonal projections instead of random reference selection. We first use an orthogonal projection $\mathbf{P} = \mathbf{I}_M - \frac{1}{M}\mathbf{1}_M\mathbf{1}_M^T$ of size $M \times M$ onto the orthogonal complement of $\mathbf{1}_M$ to eliminate $(b^2 - \|\mathbf{x}\|^2)\mathbf{1}_M$. Sequentially, we employ a second orthogonal projection $\mathbf{P}_u$ of size $M \times M$ onto the orthogonal complement of $\mathbf{Pu}$ to cancel $-2b\mathbf{Pu}$, which is given by

$$\mathbf{P}_u = \mathbf{I}_M - \frac{\mathbf{Puu}^T\mathbf{P}}{\mathbf{u}^T\mathbf{Pu}}. \qquad (6.16)$$

Thus, premultiplying (6.3) with $\mathbf{P}_u\mathbf{P}$, we obtain

$$\mathbf{P}_u\mathbf{P}\phi = 2\mathbf{P}_u\mathbf{P}\mathbf{X}_a^T\mathbf{x} + \mathbf{P}_u\mathbf{Pm}, \qquad (6.17)$$

which is linear w.r.t. $\mathbf{x}$. The price paid for applying these two projections is the loss of information. The rank of $\mathbf{P}_u\mathbf{P}$ is $M - 2$, which means that $M \geq l + 2$ still has to be fulfilled as before to obtain an unconstrained LS or WLS estimate of $\mathbf{x}$ based on (6.17). In a different way, $\mathbf{P}_u\mathbf{P}$ can be achieved directly by calculating an orthogonal projection onto the orthogonal complement of $[\mathbf{1}_M, \mathbf{u}]$. Let us define the nullspace $\mathcal{N}(\mathbf{U}^T) = \text{span}(\mathbf{1}_M, \mathbf{u})$, and $\mathcal{R}(\mathbf{U}) \oplus \mathcal{N}(\mathbf{U}^T) = \mathbb{R}^M$, where $\mathcal{R}(\mathbf{U})$ is the column space of $\mathbf{U}$, $\oplus$ denotes the direct sum of two linearly independent subspaces and $\mathbb{R}^M$ is the M-dimensional vector space. Therefore, $\mathbf{P}_u\mathbf{P}$ is the projection onto $\mathcal{R}(\mathbf{U})$. Note that the rank of $\mathbf{P}_u\mathbf{P}\mathbf{X}_a^T$ has to be

equal to l, which indicates that the anchors should not be co-linear for both 2-D and 3-D or co-planar for 3-D. A special case occurs when $\mathbf{u} = k\mathbf{1}_M$, where k is any positive real number. In this case, $\mathbf{P}$ can cancel out both $(b^2 - \|\mathbf{x}\|^2)\mathbf{1}_M$ and $-2b\mathbf{u}$, and one projection is enough, leading to the condition $M \geq l + 1$. The drawback though is that we can then only estimate $\mathbf{x}$ and $b^2 - \|\mathbf{x}\|^2 - 2bk$ due to the dependence between $\mathbf{u}$ and $\mathbf{1}_M$ according to (6.3). The SM method indicates all the insights mentioned above, which cannot be easily observed by the unconstrained estimators.

Based on (6.17), the LS and WLS estimate of $\mathbf{x}$ is respectively given by,

$$\hat{\mathbf{x}} = \frac{1}{2}(\mathbf{X}_a\mathbf{P}\mathbf{P}_u\mathbf{P}\mathbf{X}_a^T)^{-1}\mathbf{X}_a\mathbf{P}\mathbf{P}_u\mathbf{P}\phi, \tag{6.18}$$

and

$$\hat{\mathbf{x}} = \frac{1}{2}(\mathbf{X}_a\mathbf{Q}\mathbf{X}_a^T)^{-1}\mathbf{X}_a\mathbf{Q}\phi, \tag{6.19}$$

where $\mathbf{Q}$ is an aggregate weighting matrix of size $M \times M$. The optimal $\mathbf{Q}$ is given by

$$\mathbf{Q}^* = \mathbf{P}\mathbf{P}_u(\mathbf{P}_u\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}\mathbf{P}_u)^\dagger\mathbf{P}_u\mathbf{P} \tag{6.20}$$

$$= (\mathbf{P}_u\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}\mathbf{P}_u)^\dagger, \tag{6.21}$$

where the pseudo-inverse ($\dagger$) is employed, because the argument is rank deficient. Note that $\mathbf{P}_u\mathbf{P}$ is the projection onto $\mathcal{R}(\mathbf{U})$, and is applied to both sides of $\boldsymbol{\Sigma}$. Thus, $(\mathbf{P}_u\mathbf{P}\boldsymbol{\Sigma}\mathbf{P}\mathbf{P}_u)^\dagger$ is still in $\mathcal{R}(\mathbf{U})$, and would not change with applying the projection again. As a result, we can simplify (6.20) as (6.21). Consequently, $\mathbf{Q}^*$ is the pseudo-inverse of the matrix obtained by projecting the columns and rows of $\boldsymbol{\Sigma}$ onto $\mathcal{R}(\mathbf{U})$, which is of rank $M - 2$. We remark that $\hat{\mathbf{x}}$ in (6.18) (or (6.19)) is identical to the one in (6.7) (or (6.8)) according to [123]. The SM method and the unconstrained LS (or WLS) method lead to the same result. Therefore, $\hat{\mathbf{x}}$ in (6.18) and (6.7) (or in (6.19) and (6.8)) are all REFF-USR-LS (or REFF-USR-WLS) estimates.

Revisiting existing localization algorithms

As we mentioned before, all the REFB-USR-LS methods suffer from a poor reference selection. There are some efforts to improve the reference selection [87, 118, 119]. In [118], the operation employed to cancel $\|\mathbf{x}\|^2\mathbf{1}_M$ is equivalent to the orthogonal projection $\mathbf{P}$. All anchors are chosen as a reference once in [119] in order to obtain $M(M - 1)/2$ equations in total. A reference anchor is chosen based on the criterion of the shortest anchor-target distance measurement in [87]. However, reference-free methods are better than these heuristic reference-based methods in the sense that they cancel nuisance parameters in a systematic way. To clarify the relations between the REFB-USR and the REFF-USR approaches,

we generalize the reference selection of all the reference-based methods as a linear transformation, which is used to cancel nuisance parameters, similarly as an orthogonal projection. To eliminate $(b^2 - \|\mathbf{x}\|^2)\mathbf{1}_M$, the ith anchor is chosen as a reference to make differences. As a result, the corresponding linear transformation $\mathbf{T}_i$ of size $(M-1)\times M$ can be obtained by inserting the column vector $-\mathbf{1}_{M-1}$ after the $(i-1)$th column of $\mathbf{I}_{M-1}$, which fulfills $\mathbf{T}_i\mathbf{1}_M = \mathbf{0}_{M-1}, i \in \{1,\ldots,M\}$. For example, if the first anchor is chosen as a reference, then $\mathbf{T}_1 = [-\mathbf{1}_{M-1}, \ \mathbf{I}_{M-1}]$. Furthermore, we can write $\mathbf{T}_i\mathbf{d} = \mathbf{T}_{i1}\mathbf{d} - d_i\mathbf{1}_{M-1}$, where $\mathbf{T}_{i1}$ is achieved by replacing the ith column of $\mathbf{T}_i$ with the column vector $\mathbf{0}_{M-1}$. Applying $\mathbf{T}_i$ to both sides of (6.3), we arrive at

$$\mathbf{T}_i\boldsymbol{\phi} = 2\mathbf{T}_i\mathbf{X}_a^T\mathbf{x} - 2b\mathbf{T}_i\mathbf{u} + \mathbf{T}_i\mathbf{m}. \qquad (6.22)$$

Sequentially, we investigate the second linear transformation $\mathbf{M}_j$ of size $(M-2)\times (M-1)$, which fulfills $\mathbf{M}_j\mathbf{T}_i\mathbf{u} = \mathbf{0}_{M-2}, j \in \{1,\ldots,M\}$ and $j \neq i$. As a result, the nullspace $\mathcal{N}(\mathbf{M}_j\mathbf{T}_i) = \mathrm{span}(\mathbf{1}_M, \mathbf{u}) = \mathcal{N}(\mathbf{U}^T)$, and $\mathcal{R}(\mathbf{T}_i^T\mathbf{M}_j^T) = \mathcal{R}(\mathbf{U})$. Note that $b = 0$ in [9,87,118,119,123,126], which means that there is no need to apply $\mathbf{M}_j$ in these works. But the double differencing method in [117] is equivalent to employing $\mathbf{M}_j$, and thus the results of [117] can be used to design $\mathbf{M}_j$. Let us first define a matrix $\bar{\mathbf{T}}_{j1}$ of size $(M-2)\times (M-1)$ similarly as $\mathbf{T}_{i1}$ using the column vector $\mathbf{0}_{M-2}$ instead of $\mathbf{0}_{M-1}$. When the jth anchor is chosen as a reference and $j < i$, $\mathbf{M}_j$ can be obtained by inserting the column vector $-(1/(u_j - u_i))\mathbf{1}_{M-2}$ after the $(j-1)$th column of the matrix $\mathrm{diag}(\bar{\mathbf{T}}_{j1}(\mathbf{1}_{M-1} \oslash (\mathbf{T}_i\mathbf{u})))$, where $\oslash$ is element-wise division. If $j > i$, then $\mathbf{M}_j$ can be obtained by inserting the column vector $-(1/(u_j - u_i))\mathbf{1}_{M-2}$ after the $(j-2)$th column of the matrix $\mathrm{diag}(\bar{\mathbf{T}}_{(j-1)1}(\mathbf{1}_{M-1} \oslash (\mathbf{T}_i\mathbf{u})))$. For example, if the first anchor is chosen to cancel out $(b^2 - \|\mathbf{x}\|^2)\mathbf{1}_M$ ($\mathbf{T}_1$ is used), and the second anchor is chosen to eliminate $\mathbf{T}_1\mathbf{u}$, then $\mathbf{M}_2$ is given by

$$\mathbf{M}_2 = \begin{bmatrix} -1/(u_2 - u_1) & 1/(u_3 - u_1) & & & \\ -1/(u_2 - u_1) & & 1/(u_4 - u_1) & & \\ \vdots & & & \ddots & \\ -1/(u_2 - u_1) & & & & 1/(u_M - u_1) \end{bmatrix}. \qquad (6.23)$$

Premultiplying $\mathbf{M}_j\mathbf{T}_i$ to both sides of (6.3), we achieve

$$\mathbf{M}_j\mathbf{T}_i\boldsymbol{\phi} = 2\mathbf{M}_j\mathbf{T}_i\mathbf{X}_a^T\mathbf{x} + \mathbf{M}_j\mathbf{T}_i\mathbf{m}. \qquad (6.24)$$

Consequently, the general form of the REFB-USR-LS and the REFB-USR-WLS estimates are derived in the same way as (6.18) and (6.19) by replacing $\mathbf{P}\mathbf{P}_u\mathbf{P}$ and $\mathbf{Q}$ with $\mathbf{T}_i^T\mathbf{M}_j^T\mathbf{M}_j\mathbf{T}_i$ and $\mathbf{Q}_{i,j}$, respectively. We do not repeat these equations for the sake of brevity. Note that $\mathbf{Q}_{i,j}$ is an aggregate weighting matrix of size

$M \times M$. The optimal $\mathbf{Q}_{i,j}$ is given by

$$
\begin{aligned}
\mathbf{Q}_{i,j}^* &= \mathbf{T}_i^T \mathbf{M}_j^T (\mathbf{M}_j \mathbf{T}_i \boldsymbol{\Sigma} \mathbf{T}_i^T \mathbf{M}_j^T)^{-1} \mathbf{M}_j \mathbf{T}_i & (6.25) \\
&= \left[(\mathbf{M}_j \mathbf{T}_i)^\dagger \mathbf{M}_j \mathbf{T}_i \boldsymbol{\Sigma} \mathbf{T}_i^T \mathbf{M}_j^T (\mathbf{T}_i^T \mathbf{M}_j^T)^\dagger \right]^\dagger, & (6.26)
\end{aligned}
$$

where $(\mathbf{M}_j \mathbf{T}_i)^\dagger \mathbf{M}_j \mathbf{T}_i = \mathbf{T}_i^T \mathbf{M}_j^T (\mathbf{T}_i^T \mathbf{M}_j^T)^\dagger = \mathbf{T}_i^T \mathbf{M}_j^T (\mathbf{M}_j \mathbf{T}_i \mathbf{T}_i^T \mathbf{M}_j^T)^{-1} \mathbf{M}_j \mathbf{T}_i$, which is also the projection onto $\mathcal{R}(\mathbf{U})$, and thus is equivalent to $\mathbf{P}_u \mathbf{P}$. The equality between (6.25) and (6.26) can be verified using a property of the pseudo-inverse[†]. Hence, $\mathbf{Q}_{i,j}^*$ is of rank $M - 2$, and $\mathbf{Q}_{i,j}^* = \mathbf{Q}^*, i, j \in \{1, \dots, M\}$ with $i \neq j$. As a result, the REFB-USR-WLS estimate and the REFF-USR-WLS estimate are identical if the optimal weighting matrix is used. Hence, the optimal weighting matrix can compensate the impact of random reference selection. However, since $\boldsymbol{\Sigma}$ depends on the unknown $\mathbf{d}$, the optimal weighting matrix can only be approximated iteratively. Also note that the REFB-USR-LS estimate suffers from the ad-hoc reference selection, while the REFF-USR-LS estimate is independent of the reference selection.

6.2.3 Localization Based on Squared Differences of TOA measurements

Proposed localization algorithms

Let us recall (6.1) here, i.e.,

$$
\mathbf{u} - b\mathbf{1}_M = \mathbf{d} + \mathbf{n}. \tag{6.27}
$$

In general, b is regarded as a nuisance parameter. Instead of first carrying out element-wise multiplication at both sides of (6.27), we can also try to get rid of b before element-wise multiplication. By choosing a reference anchor, and then subtracting the TOAs of other anchors from the TOA of the reference anchor [9], $M - 1$ TDOAs are obtained and b is canceled out. Note that these TDOAs are achieved differently from the TDOAs obtained directly by cross-correlating the received signals from different anchors. The obvious drawback of this conventional scheme is again the reference dependency. On the other hand, since b is a common term in (6.1), we can again apply $\mathbf{P}$ to eliminate $-b\mathbf{1}_M$ instead of randomly choosing a reference anchor. Then we arrive at

$$
\mathbf{P}\mathbf{u} = \mathbf{P}\mathbf{d} + \mathbf{P}\mathbf{n}. \tag{6.28}
$$

Note that $\mathbf{P}\mathbf{u} = \mathbf{u} - \bar{u}\mathbf{1}_M$, where $\bar{u}$ is the average TOA. Thus, $\mathbf{P}\mathbf{u}$ represents the differences between the anchor TOAs and the average TOA. Moreover, $\mathbf{P}\mathbf{d} = \mathbf{d} - \bar{d}\mathbf{1}_M$, where $\bar{d} = \frac{1}{M} \sum_{i=1}^{M} d_i$ is the unknown average of the distances between

---

[†]Given the matrix $\mathbf{C}$ of size $n \times r$ and the matrix $\mathbf{D}$ of size $r \times m$ both of rank r, then if $\mathbf{A} = \mathbf{CD}$, it holds that $\mathbf{A}^\dagger = \mathbf{D}^\dagger \mathbf{C}^\dagger$ [130].

the target node and the anchors, and $\mathbf{Pn} = \mathbf{n} - \bar{n}\mathbf{1}_M$, where $\bar{n} = \frac{1}{M}\sum_{i=1}^{M} n_i$. Thus, (6.28) can be rewritten as

$$\mathbf{Pu} + (\bar{d} + \bar{n})\mathbf{1}_M = \mathbf{d} + \mathbf{n}, \tag{6.29}$$

By making element-wise multiplication of (6.29) and re-arranging all the terms, we achieve

$$\begin{aligned}
&\boldsymbol{\psi}_a - (\mathbf{Pu}) \odot (\mathbf{Pu}) \\
= \quad &2\mathbf{X}_a^T\mathbf{x} + 2\bar{d}\mathbf{Pu} + (\bar{d}^2 - \|\mathbf{x}\|^2)\mathbf{1}_M + \mathbf{m} + \bar{n}^2\mathbf{1}_M + 2\bar{n}(\bar{d}\mathbf{1}_M + \mathbf{Pu})
\end{aligned} \tag{6.30}$$

where $\boldsymbol{\psi}_a = [\|\mathbf{x}_1\|^2, \|\mathbf{x}_2\|^2, \ldots, \|\mathbf{x}_M\|^2]^T$ and $\mathbf{m} = -(2\mathbf{d} \odot \mathbf{n} + \mathbf{n} \odot \mathbf{n})$ as before. Using the SM method to obtain an unconstrained LS estimate of $\mathbf{x}$ alone, we employ again two projections $\mathbf{P}$ and $\mathbf{P}_u$, and arrive at

$$\mathbf{P}_u\mathbf{P}(\boldsymbol{\psi}_a - (\mathbf{Pu}) \odot (\mathbf{Pu})) = 2\mathbf{P}_u\mathbf{P}\mathbf{X}_a^T\mathbf{x} + \mathbf{P}_u\mathbf{Pm}, \tag{6.31}$$

the right hand side of which is exactly the same as the one in (6.17), and thus we can state $\mathbf{P}_u\mathbf{P}(\boldsymbol{\psi}_a - (\mathbf{Pu}) \odot (\mathbf{Pu})) = \mathbf{P}_u\mathbf{P}\boldsymbol{\phi}$. Note that although (6.30) is different from (6.3), we find that (6.31) and (6.17) become equivalent after premultiplying $\mathbf{P}_u\mathbf{P}$. Furthermore, $(\mathbf{Pu}) \odot (\mathbf{Pu})$ can be labeled as a squared range difference (SRD) term. As a result, the unconstrained LS and WLS estimate of $\mathbf{x}$ based on (6.31), which are named the reference-free USRD-LS (REFF-USRD-LS) estimate and the REFF-USRD-WLS estimate, are exactly the same as the REFF-USR-LS estimate (6.18) and the REFF-USR-WLS estimate (6.19), respectively. We do not repeat them here in the interest of brevity. Moreover, the constrained LS and WLS based on (6.30), namely the REFF-SRD-CLS estimate and the REFF-SRD-CWLS estimate, are identical to the REFF-SR-CLS and the REFF-SR-CWLS estimate (6.14) as well.

Revisiting existing localization algorithms

Existing methods choose a reference anchor to obtain range differences, and further investigate low-complexity closed-form LS or WLS solutions. Thus, we call them reference-based USRD-LS (REFB-USRD-LS) and REFB-USRD-WLS approaches. To expose interesting links among the different reference-based or reference-free squared-range-based or squared-range-difference-based approaches, we generalize the conventional REFB-USRD-LS and REFB-USRD-WLS approaches [9] in the same way as in Section 6.2.2. The reference selection can be generalized by a linear transformation similarly as in Section 6.2.2. In order to eliminate $-b\mathbf{1}_M$ in (6.27), the ith anchor is chosen as a reference, thus $\mathbf{T}_i$ defined in Section 6.2.2 is employed, which fulfills $\mathbf{T}_i\mathbf{1}_M = \mathbf{0}_{M-1}$. Applying $\mathbf{T}_i$ instead of $\mathbf{P}$ to (6.27), following the same operations to obtain (6.30), and noting that

$(\mathbf{T}_{i1}(\mathbf{d}+\mathbf{n}))\odot(\mathbf{T}_{i1}(\mathbf{d}+\mathbf{n})) = \mathbf{T}_{i1}((\mathbf{d}+\mathbf{n})\odot(\mathbf{d}+\mathbf{n}))$, we arrive at

$$\mathbf{T}_i\boldsymbol{\psi}_a - (\mathbf{T}_i\mathbf{u})\odot(\mathbf{T}_i\mathbf{u}) = 2\mathbf{T}_i\mathbf{X}_a^T\mathbf{x} + 2d_i\mathbf{T}_i\mathbf{u} + \mathbf{T}_i\mathbf{m} + 2n_i\mathbf{T}_i\mathbf{u}, \quad (6.32)$$

which is different from (6.30), and has only one nuisance parameter d_i at the right hand side. Ignoring the relation between $\mathbf{x}$ and d_i, we still have two ways to deal with d_i. The first one is to estimate $\mathbf{x}$ and d_i together [123], which means we only use a reference once for calculating the TDOAs. The second one is again to apply $\mathbf{M}_j$, which fulfills $\mathbf{M}_j\mathbf{T}_i\mathbf{u} = \mathbf{0}_{M-2}$. It employs two different references, one for calculating the TDOAs, and the other for eliminating the nuisance parameter. In order to distinguish these two, we call them the REFB-USRD-LS(1) and the REFB-USRD-LS(2) estimate, respectively, where the number between brackets indicates the number of references used in the approach. In the same way as we clarified the equivalence between the REFF-USRD-LS and the REFF-USR-LS estimate in the previous subsection, we can easily confirm the equivalence between the REFB-USRD-LS(2) (or the REFB-USRD-WLS(2)) and the REFB-USR-LS (or the REFB-USR-WLS) estimate of Section 6.2.2. We omit the details for the sake of brevity. Furthermore, we recall that similarly as above we could have dealt with $-2b\mathbf{T}_i\mathbf{u}$ in (6.22) in two different ways. But since $b = 0$ in [9,87,118,119,123,126], there are no discussions about these two different ways in literature, and we do not distinguish between them in the REFB-USR-LS method.

Since there is no counterpart of the REFB-USRD-LS(1) estimate in Section 6.2.2 for the squared-range-based methods, we briefly discuss the REFB-USRD-LS(1) estimate to complete the investigation of the links among all the estimators based on TOA measurements. Employing the SM method, we again use an orthogonal projection $\mathbf{P}_i$ of size $(M-1)\times(M-1)$ onto the orthogonal complement of $\mathbf{T}_i\mathbf{u}$ to fulfill $\mathbf{P}_i\mathbf{T}_i\mathbf{u} = \mathbf{0}_{M-1}$, which can be derived in the same way as (6.16) by replacing $\mathbf{I}_M$ and $\mathbf{Pu}$ with $\mathbf{I}_{M-1}$ and $\mathbf{T}_i\mathbf{u}$, respectively. As a result, $\mathcal{N}(\mathbf{P}_i\mathbf{T}_i) = \text{span}(\mathbf{1}_M,\mathbf{u}) = \mathcal{N}(\mathbf{U}^T)$ and $\mathcal{R}(\mathbf{T}_i^T\mathbf{P}_i) = \mathcal{R}(\mathbf{U})$. Premultiplying (6.32) with $\mathbf{P}_i$, we obtain

$$\mathbf{P}_i\mathbf{T}_i\boldsymbol{\psi}_a - \mathbf{P}_i((\mathbf{T}_i\mathbf{u})\odot(\mathbf{T}_i\mathbf{u})) = 2\mathbf{P}_i\mathbf{T}_i\mathbf{X}_a^T\mathbf{x} + \mathbf{P}_i\mathbf{T}_i\mathbf{m}. \quad (6.33)$$

Note that $\mathbf{P}_i((\mathbf{T}_i\mathbf{u})\odot(\mathbf{T}_i\mathbf{u})) = \mathbf{P}_i\mathbf{T}_i(\mathbf{u}\odot\mathbf{u})$ (see Appendix 6.B for a proof), and thus we can state $\mathbf{P}_i\mathbf{T}_i\boldsymbol{\psi}_a - \mathbf{P}_i((\mathbf{T}_i\mathbf{u})\odot(\mathbf{T}_i\mathbf{u})) = \mathbf{P}_i\mathbf{T}_i\boldsymbol{\phi}$. Consequently, the REFB-USRD-LS(1) and the REFB-USRD-WLS(1) estimates can also be written as (6.18) and (6.19) by replacing $\mathbf{PP}_u\mathbf{P}$ and $\mathbf{Q}$ with $\mathbf{T}_i^T\mathbf{P}_i\mathbf{T}_i$ and $\mathbf{Q}_i$, respectively. We do not repeat the equations in the interest of brevity. We remark that $\mathbf{Q}_i$ is again an aggregate weighting matrix of size $M\times M$, and the optimal $\mathbf{Q}_i$ of rank $(M-2)$ is given by

$$\mathbf{Q}_i^* = \mathbf{T}_i^T\mathbf{P}_i(\mathbf{P}_i\mathbf{T}_i\boldsymbol{\Sigma}\mathbf{T}_i^T\mathbf{P}_i)^\dagger\mathbf{P}_i\mathbf{T}_i \quad (6.34)$$

$$= (\mathbf{V}_i\mathbf{V}_i^T\boldsymbol{\Sigma}\mathbf{V}_i\mathbf{V}_i^T)^\dagger, \quad (6.35)$$

where $\mathbf{V}_i$ is of size $M \times (M-2)$, and collects the right singular vectors corresponding to the $M-2$ nonzero singular values of $\mathbf{P}_i\mathbf{T}_i$. We derive (6.35) in Appendix 6.C, and prove that $\mathbf{V}_i\mathbf{V}_i^T$ is the projection onto $\mathcal{R}(\mathbf{U})$. As a result, $\mathbf{Q}_i^* = \mathbf{Q}_{i,j}^* = \mathbf{Q}^*, i,j \in \{1,\ldots,M\}$ and $i \neq j$.

Based on the above discussions, we achieve the important conclusion that the REFF-USRD-WLS, the REFB-USRD-WLS(1), the REFB-USRD-WLS(2), the REFF-USR-WLS, and the REFB-USR-WLS estimate are all identical if the optimal weighting matrix is adopted. The optimal weighting matrix releases the reference-based methods from the influence of a random reference selection. Moreover, the REFF-USR-LS and the REFF-USRD-LS estimate are identical, and free from a reference selection, whereas the REFB-USR-LS and the REFB-USRD-LS(2) estimate are equivalent, but still suffer from a poor reference selection.

To further improve the localization accuracy, a constrained WLS estimate based on (6.32) can be pursued considering the relation between $\mathbf{x}$ and d_i similarly as in [126]. We call it the reference-based SRD CWLS (REFB-SRD-CWLS) estimate. Denoting $\mathbf{z} = [\mathbf{x}^T, d_i]^T$, $\mathbf{B}_i = 2\mathbf{T}_i[\mathbf{X}_a^T, \mathbf{u}]$ and $\boldsymbol{\varrho}_i = \mathbf{T}_i\boldsymbol{\psi}_a - (\mathbf{T}_i\mathbf{u}) \odot (\mathbf{T}_i\mathbf{u})$, it is given by,

$$\hat{\mathbf{z}} = \arg\min_{\hat{\mathbf{z}}}(\boldsymbol{\varrho}_i - \mathbf{B}_i\mathbf{z})^T\mathbf{W}_i(\boldsymbol{\varrho}_i - \mathbf{B}_i\mathbf{z}) \tag{6.36}$$

subject to

$$(\mathbf{z} - \mathbf{z}_i)^T\mathbf{L}(\mathbf{z} - \mathbf{z}_i) = 0 \quad \text{and} \quad [\mathbf{z}]_{l+1} \geq 0, \tag{6.37}$$

where $\mathbf{W}_i$ is a weighting matrix of size $(M-1) \times (M-1)$, $\mathbf{z}_i = [\mathbf{x}_i^T \quad 0]^T$ and

$$\mathbf{L} = \begin{bmatrix} \mathbf{I}_l & \mathbf{0}_l \\ \mathbf{0}_l^T & -1 \end{bmatrix}. \tag{6.38}$$

The method to solve this CWLS problem is proposed in [126]. We do not review it for the sake of brevity. Note that there are two constraints for (6.36) compared to one for (6.10), thus the method to solve (6.36) is different from the one to solve (6.10).

All the estimators based on TOA measurements are summarized in Tables 6.1 - 6.3. They are characterized by the number of references, the reference dependency, the minimum number of anchors, and the optimal weighting matrices. We also shed light on their relations and categorize the existing methods from literature. We remark that the authors in [131] claim that the error covariance of the optimal position estimate using TOAs with a distance bias is equivalent to the one using TDOAs regardless of the reference selection, where the error covariance is defined as the product of the position dilution of precision (PDOP) and a composite user-equivalent range error (UERE). However, a more appropriate indication of the localization performance is the Cramér-Rao bound (CRB), which is a bound

	REFF -USR-LS	REFB -USR-LS	REFF -USRD-LS	REFB -USRD-LS(1)	REFB -USRD-LS(2)
Relations	The REFF-USR-LS and the REFF-USRD-LS estimate are identical. The REFB-USR-LS and the REFB-USRD-LS(2) estimate are identical.				
no. of references	0	2	0	1	2
reference dependency	No	Yes	No	Yes	Yes
literature	Proposed	$[7, 87, 119]$	Proposed	$[9, 123]$	$[117]$
min. no. of anchors, $\mathbf{x}$ of length l	$l + 2$				

Table 6.1: LS estimators based on TOAs for locating an asynchronous target.

	REFF -USR-WLS	REFB -USR-WLS	REFF -USRD-WLS	REFB -USRD-WLS(1)	REFB -USRD-WLS(2)
Relations	The REFB-USR-WLS and the REFB-USRD-WLS(2) estimate are identical. They are all identical with optimal weighting matrices $\mathbf{Q}^* = \mathbf{Q}^*_{i,j} = \mathbf{Q}^*_i$.				
no. of references	0	2	0	1	2
reference dependency	No	Yes, with $\mathbf{Q}_{i,j}$ No, with $\mathbf{Q}^*_{i,j}$	No	Yes, with $\mathbf{Q}_i$ No, with $\mathbf{Q}^*_i$	Yes, with $\mathbf{Q}_{i,j}$ No, with $\mathbf{Q}^*_{i,j}$
literature	Proposed		Proposed		
min. no. of anchors, $\mathbf{x}$ of length l	$l + 2$				

Table 6.2: WLS estimators based on TOAs for locating an asynchronous target.

	REFF-SR-CWLS	REFF-SRD-CWLS	REFB-SRD-CWLS
Equations	(6.14)		(6.36)
no. of references	0	0	1
reference dependency	No	No	Yes, with $\mathbf{W}_i$ No, with $\mathbf{W}^*_i$
literature	Proposed	Proposed	$[126]$
min. no. of anchors, $\mathbf{x}$ of length l	$l + 2$		

Table 6.3: CLS estimators based on TOAs for locating an asynchronous target.

for unbiased estimators. Therefore, the CRB based on (6.1) for TOAs with a distance bias is derived in Appendix 6.D. Since the TDOAs in Section 6.2.3 are calculated by making differences of the TOAs in (6.1), the CRB based on these TDOAs is the same as the one based on (6.1).

6.3 Localization based on TDOA measurements

6.3.1 System Model

Let us now focus on TDOA measurements. In passive sensor array or microphone array localization, TDOA measurements are obtained directly by cross-correlating a pair of received signals. Thus, no correlation template is needed, and the clock-

offset can be canceled out immediately. We reemphasize that these TDOA measurements are different from the TDOAs calculated by subtracting the TOAs. The data model for these TDOA measurements is given by [132]

$$r_{i,j} = d_j - d_i + n_{i,j}, \ \ i,j \in \{1,2,\dots,M\}, \ \ i \neq j, \tag{6.39}$$

where $r_{i,j}$ is the TDOA measurement, which is obtained by cross-correlating the received signal from the jth anchor with the one from the ith anchor. Note that the stochastic properties of the noise terms $n_{i,j}$ are totally different from the ones of the noise terms n_i of (6.1). We approximate $n_{i,j}$ as zero-mean random variables, where $\mathrm{cov}(n_{i,j}, n_{p,q}) = E[(n_{i,j} - E[n_{i,j}])(n_{p,q} - E[n_{p,q}])] = E[n_{i,j}n_{p,q}], \ i,j,p,q, \in \{1,2,\dots,M\}, \ i \neq j,$ and $p \neq q$. Defining $\mathbf{r}_i$ as the collection of the corresponding distances to the $M-1$ TDOA measurements using the ith anchor as a reference, $\mathbf{r}_i = [r_{i,1},\dots,r_{i,i-1},r_{i,i+1},\dots,r_{i,M}]^T$, and $\mathbf{n}_i = [n_{i,1},\dots,n_{i,i-1},n_{i,i+1},\dots,n_{i,M}]^T$ as the related noise vector, we write (6.39) in vector form as

$$\mathbf{r}_i \ = \ \mathbf{T}_{i1}\mathbf{d} - d_i\mathbf{1}_{M-1} + \mathbf{n}_i. \tag{6.40}$$

Moving $-d_i\mathbf{1}_{M-1}$ to the other side, making an element-wise multiplication and re-arranging, we achieve

$$\boldsymbol{\varphi}_i \ = \ 2\mathbf{T}_i\mathbf{X}_a^T\mathbf{x} + 2d_i\mathbf{r}_i + \mathbf{m}_i, \tag{6.41}$$

where $\boldsymbol{\varphi}_i = \mathbf{T}_i\boldsymbol{\psi}_a - \mathbf{r}_i \odot \mathbf{r}_i$ and $\mathbf{m}_i = -(2(\mathbf{T}_{i1}\mathbf{d}) \odot \mathbf{n}_i + \mathbf{n}_i \odot \mathbf{n}_i)$. The stochastic properties of $\mathbf{m}_i$ are as follows

$$E[[\mathbf{m}_i]_k] \ = \ -E[[\mathbf{n}_i]_k \odot [\mathbf{n}_i]_k] \approx 0, \tag{6.42}$$

$$[\boldsymbol{\Sigma}_i]_{k,l} \ = \ E[[\mathbf{m}_i]_k[\mathbf{m}_i]_l] - E[[\mathbf{m}_i]_k]E[[\mathbf{m}_i]_l]$$

$$\approx \begin{cases} 4d_k d_l E[n_{i,k}n_{i,l}], & k < i \text{ and } l < i \\ 4d_{k+1} d_{l+1} E[n_{i,k+1}n_{i,l+1}], & k \geq i \text{ and } l \geq i \\ 4d_k d_{l+1} E[n_{i,k}n_{i,l+1}], & k < i \text{ and } l \geq i \\ 4d_{k+1} d_l E[n_{i,k+1}n_{i,l}], & k \geq i \text{ and } l < i \end{cases}, \tag{6.43}$$

where we ignore the higher order noise terms to obtain (6.43) and assume that the noise mean $E[[\mathbf{m}_i]_k] \approx 0$ under the condition of sufficiently small measurement errors. Note that the noise covariance matrix $\boldsymbol{\Sigma}_i$ of size $(M-1)\times(M-1)$ depends on the unknown $\mathbf{d}$ as well.

6.3.2 Localization Based on Squared TDOA measurements

We do not propose any new algorithms in this section, but summarize existing localization algorithms spread over different research areas and shed light on their relations. All these algorithms are categorized as reference-based SRD approaches.

Note that (6.41) looks similar to (6.32). Only the available data and the noise characteristics are different, which leads to totally different relations among the estimators as we will show in the following paragraphs. The approach to achieve the REFB-USRD-LS(1) estimate, the REFB-USRD-LS(2) estimate and the REFB-SRD-CWLS estimate (6.36) based on TOA measurements in Section 6.2.3 can be adopted here as well. The orthogonal projection $\widetilde{\mathbf{P}}_i$ of size $(M-1) \times (M-1)$ onto the complement of $\mathbf{r}_i$ is employed, which is given by (6.16), where we replace $\mathbf{I}_M$ and $\mathbf{Pu}$ with $\mathbf{I}_{M-1}$ and $\mathbf{r}_i$. Let us define the nullspace $\mathcal{N}(\widetilde{\mathbf{U}}_i^T) = \mathrm{span}(\mathbf{r}_i)$, and $\mathcal{R}(\widetilde{\mathbf{U}}_i) \oplus \mathcal{N}(\widetilde{\mathbf{U}}_i^T) = \mathbb{R}^{M-1}$. Therefore, $\widetilde{\mathbf{P}}_i$ is the projection onto $\mathcal{R}(\widetilde{\mathbf{U}}_i)$. As a result, the REFB-USRD-LS(1) and REFB-USRD-WLS(1) estimate based on TDOA measurements is respectively given by,

$$\hat{\mathbf{x}} = -\frac{1}{2}(\mathbf{X}_a \mathbf{T}_i^T \widetilde{\mathbf{P}}_i \mathbf{T}_i \mathbf{X}_a^T)^{-1} \mathbf{X}_a \mathbf{T}_i^T \widetilde{\mathbf{P}}_i \boldsymbol{\varphi}_i, \qquad (6.44)$$

and

$$\hat{\mathbf{x}} = -\frac{1}{2}(\mathbf{X}_a \mathbf{T}_i^T \widetilde{\mathbf{Q}}_i \mathbf{T}_i \mathbf{X}_a^T)^{-1} \mathbf{X}_a \mathbf{T}_i^T \widetilde{\mathbf{Q}}_i \boldsymbol{\varphi}_i, \qquad (6.45)$$

where $\widetilde{\mathbf{Q}}_i$ is an aggregate weighting matrix of size $(M-1) \times (M-1)$ as well. Note that (6.44) (or (6.45)) differs from (6.18) (or (6.19)) since $M-1$ TDOA measurements are used instead of M TOA measurements. The optimal $\widetilde{\mathbf{Q}}_i$ is given by

$$\widetilde{\mathbf{Q}}_i^* = \widetilde{\mathbf{P}}_i (\widetilde{\mathbf{P}}_i \boldsymbol{\Sigma}_i \widetilde{\mathbf{P}}_i)^\dagger \widetilde{\mathbf{P}}_i \qquad (6.46)$$

$$= (\widetilde{\mathbf{P}}_i \boldsymbol{\Sigma}_i \widetilde{\mathbf{P}}_i)^\dagger, \qquad (6.47)$$

where $\widetilde{\mathbf{Q}}_i^*, i \in \{1, \ldots, M\}$ is the pseudo-inverse of the matrix achieved by projecting the columns and rows of $\boldsymbol{\Sigma}_i$ onto $\mathcal{R}(\widetilde{\mathbf{U}}_i)$, which is of rank $M-2$. We remark that the REFB-USRD-LS(1) estimate (6.44) is equivalent to the ones in [123,133].

Let us also revisit the REFB-USRD-LS(2) estimate and the REFB-USRD-WLS(2) estimate based on TDOA measurements. A linear transformation $\widetilde{\mathbf{M}}_j$ of size $(M-2) \times (M-1)$, which fulfills $\widetilde{\mathbf{M}}_j \mathbf{r}_i = \mathbf{0}_{M-2}$, can be devised in the same way as $\mathbf{M}_j$ by replacing $\mathbf{T}_i \mathbf{u}$ and $1/(u_j - u_i)$ with $\mathbf{r}_i$ and $1/r_{i,j}$, respectively. Thus, $\mathcal{R}(\widetilde{\mathbf{M}}_j^T) = \mathcal{R}(\widetilde{\mathbf{U}}_i)$. Note that another heuristic method to obtain $\widetilde{\mathbf{M}}_j$ is proposed in [121]. As a result, the general form of the REFB-USRD-LS(2) and the REFB-USRD-WLS(2) estimates can be derived in the same way as (6.44) and (6.45) by replacing $\widetilde{\mathbf{P}}_i$ and $\widetilde{\mathbf{Q}}_i$ with $\widetilde{\mathbf{M}}_j^T \widetilde{\mathbf{M}}_j$ and $\widetilde{\mathbf{Q}}_{i,j}$, respectively. Note that $\widetilde{\mathbf{Q}}_{i,j}$ is also an aggregate weighting matrix of size $(M-1) \times (M-1)$. The optimal $\widetilde{\mathbf{Q}}_{i,j}^*$

is given by

$$\widetilde{\mathbf{Q}}_{i,j}^* = \widetilde{\mathbf{M}}_j^T (\widetilde{\mathbf{M}}_j \mathbf{\Sigma}_i \widetilde{\mathbf{M}}_j^T)^{-1} \widetilde{\mathbf{M}}_j \qquad (6.48)$$

$$= \left[(\widetilde{\mathbf{M}}_j)^\dagger \widetilde{\mathbf{M}}_j \mathbf{\Sigma}_i \widetilde{\mathbf{M}}_j^T (\widetilde{\mathbf{M}}_j^T)^\dagger \right]^\dagger, \qquad (6.49)$$

where $(\widetilde{\mathbf{M}}_j)^\dagger \widetilde{\mathbf{M}}_j = \widetilde{\mathbf{M}}_j^T (\widetilde{\mathbf{M}}_j^T)^\dagger = \widetilde{\mathbf{M}}_j^T (\widetilde{\mathbf{M}}_j \widetilde{\mathbf{M}}_j^T)^{-1} \widetilde{\mathbf{M}}_j$ is also the projection onto $\mathcal{R}(\widetilde{\mathbf{U}}_i)$, which means that $\widetilde{\mathbf{Q}}_{i,j}^* = \widetilde{\mathbf{Q}}_i^*, i,j \in \{1, \ldots, M\}$ and $i \neq j$. The REFB-USRD-LS(2) estimate and the REFB-USRD-WLS(2) estimate based on TDOA measurements are generalizations of the estimators proposed in [121]. However, the noise covariance matrix in [121] is a diagonal matrix, and the noise covariance matrix $\mathbf{\Sigma}_i$ here is a full matrix.

We remark here that with the optimal weighting matrix, the REFB-USRD-WLS(1) estimate (6.45) and the REFB-USRD-WLS(2) estimate based on the same set of TDOA measurements are identical. However, the optimal weighting matrix cannot decouple the reference dependency. The performance of all the estimates still depends on the reference selection, since the reference dependency is an inherent property of the available measurement data. To further improve the localization performance, the REFB-SRD-CWLS estimate based on (6.41) can be derived in the same way as the estimate (6.36) by replacing $\boldsymbol{\varrho}_i$ and $\mathbf{B}_i$ with $\boldsymbol{\varphi}_i$ and $2[\mathbf{T}_i \mathbf{X}_a^T, \mathbf{r}_i]$, respectively. A solution to this CLS problem is presented in [126].

Note that all the above estimators are based on a so-called nonredundant set of TDOA measurements [132], resulting in reference dependency. Recently, a SM method based on the full set of TDOA measurements has been proposed in [105], labeled "reference-free TDOA source localization". It is reference-free in the sense that every anchor plays the role of reference, as in [119], thus there is no need to specifically choose one. We revisit the proposed method in [105] here to clarify its relation to our framework. Let us define $\mathbf{D}_r = [\widetilde{\mathbf{r}}_1, \ \widetilde{\mathbf{r}}_1, \ \ldots, \ \widetilde{\mathbf{r}}_M]$, where $\widetilde{\mathbf{r}}_i$ can be achieved by inserting a 0 in $\mathbf{r}_i$ between $r_{i,i-1}$ and $r_{i,i+1}$. Using our notations, we can rewrite (22) of [105] as

$$\frac{1}{2M}(\mathbf{D}_r \odot \mathbf{D}_r)\mathbf{1}_M - \frac{1}{M}\mathbf{D}_r\mathbf{d} = \frac{1}{2}\mathbf{P}\boldsymbol{\psi}_a - \mathbf{P}\mathbf{X}_a^T\mathbf{x}. \qquad (6.50)$$

Then, a matrix $\mathbf{G}$ of size $(M-2) \times M$, which fulfills $\mathbf{G}\mathbf{D}_r = \mathbf{0}_{M-2}$, can be obtained by exploring the nullspace of $\mathbf{D}_r$ using the singular value decomposition (SVD). Consequently, an LS estimator of $\mathbf{x}$ is given by

$$\hat{\mathbf{x}} = \frac{1}{2}(\mathbf{X}_a \mathbf{P}\mathbf{G}^T \mathbf{G}\mathbf{P}\mathbf{X}_a^T)^{-1}\mathbf{X}_a \mathbf{P}\mathbf{G}^T \mathbf{G}(\mathbf{P}\boldsymbol{\psi}_a - \frac{1}{M}(\mathbf{D}_r \odot \mathbf{D}_r)\mathbf{1}_M). \ (6.51)$$

Note that $\mathbf{D}_r = [\mathbf{d}, \ \mathbf{1}_M]\begin{bmatrix} \mathbf{1}_M^T \\ -\mathbf{d}^T \end{bmatrix}$ without noise, and $\mathbf{G}\mathbf{D}_r = \mathbf{0}_{M-2}$. Thus, $\mathbf{1}_M$ is in the nullspace of $\mathbf{G}$. As $\mathbf{P}$ is the projection onto the orthogonal complement of $\mathbf{1}_M$, $\mathbf{G}\mathbf{P}$ is still of rank $M-2$ with probability 1. In a different way, we can

	REFB -USRD-LS(1)	REFB -USRD-WLS(1)	REFB -USRD-LS(2)	REFB -USRD-WLS(2)	REFB -SRD-CWLS
Relations	The REFB-USRD-WLS(1) and the REFB-USRD-WLS(2) estimate are identical with the optimal weighting matrices $\widetilde{\mathbf{Q}}_i^* = \widetilde{\mathbf{Q}}_{i,j}^*$.				
no. of references	1	1	2	2	1
reference dependency	Yes	Yes	Yes	Yes	Yes
literature	[120, 134] [105, 123, 133]	[122]	[121]	[121]	[6, 126]
min. no. of anchors, $\mathbf{x}$ of length l	$l + 2$				

Table 6.4: LS, WLS and CWLS estimators based on TDOAs for locating an asynchronous target.

make use of the full set of TDOA measurements similarly as the second extension of the approach proposed in [133]. We collect (6.41) in vector form as

$$
\begin{bmatrix} \boldsymbol{\varphi}_1 \\ \boldsymbol{\varphi}_2 \\ \vdots \\ \boldsymbol{\varphi}_M \end{bmatrix} = 2 \begin{bmatrix} \mathbf{T}_1 \\ \mathbf{T}_2 \\ \vdots \\ \mathbf{T}_M \end{bmatrix} \mathbf{X}_a^T \mathbf{x} + 2 \begin{bmatrix} \mathbf{r}_1 & & & \\ & \mathbf{r}_2 & & \\ & & \ddots & \\ & & & \mathbf{r}_M \end{bmatrix} \mathbf{d} + \begin{bmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \vdots \\ \mathbf{m}_M \end{bmatrix}
\tag{6.52}
$$

As a result, a LS estimator of $\mathbf{x}$ and $\mathbf{d}$ can be derived based on (6.52). We do not detail it in the interest of brevity.

Furthermore, as indicated in [132], an optimal nonredundant set can be achieved by the optimum conversion of the full TDOA set in order to approach the same localization performance, and the use of this optimal nonredundant set is recommended to reduce the complexity. Because [132] relies on the assumption that the received signals at the anchors are corrupted by noise with equal variances, the optimal nonredundant set can be estimated by a LS estimator. This is not the case here however, where it should be estimated by a WLS estimator, which requires the knowledge of the stochastic properties of the noise.

We summarize the characteristics of all the estimators based on TDOA measurements in Table 6.4. With the nonredundant TDOA measurement set of length $M - 1$, the estimator performance suffers from a poor reference selection. Although the performance improves with the full set or the optimal nonredundant set, it first has to measure the full set of TDOAs of length $M(M - 1)/2$.

6.4   Numerical Results

6.4.1   Noise statistics

In order to make a fair comparison between the localization performance of the different estimators using TOA measurements and TDOA measurements, we de-

rive the statistics of n_i and $n_{i,j}$ based on the same received signal models. The received signal is modeled by [105]

$$z_i(n) = \frac{\kappa}{d_i} s(n - \tau_i) + e_i(n), \;\; n = 0, 1, \ldots, N-1, \tag{6.53}$$

where N is the number of samples, κ is a constant parameter, $s(n)$ is the source signal, and $e_i(n)$ and τ_i are respectively the additive noise and the delay at the ith node. We assume that $s(n)$ is a zero-mean white sequence with variance σ_s^2, and $e_i(n)$ is also a zero-mean white sequence with variance σ_e^2, independent from the other noise sequences and $s(n)$.

For the TOA-based approaches, we assume knowledge of the template $s(n)$, and estimate τ_i by cross-correlating the received signal with the clean template:

$$\hat{\tau}_i = \operatorname*{argmax}_{\tau_i} \left\{ \sum_{n=0}^{N-1} z_i(n) s(n - \tau_i) \right\}. \tag{6.54}$$

Since there is an unknown bias due to asynchronous nodes, the distance u_i corresponding to the timestamp is modeled as $u_i = c\hat{\tau}_i = d_i + b + n_i$, where c is the signal propagation speed. The statistical properties of n_i can be derived in a similar way as in [132], and are given by

$$E[n_i] = 0, \tag{6.55}$$

$$\operatorname{cov}(n_i, n_j) = E[n_i n_j]$$
$$= \begin{cases} \sigma_i^2 = \dfrac{3c^2}{N\pi^2\kappa^2} \dfrac{d_i^2}{\text{SNR}} & i = j \\ 0 & i \neq j \end{cases}, \tag{6.56}$$

where $\text{SNR} = \sigma_s^2/\sigma_e^2$. We remark that in reality, it is very difficult to obtain a clean template, since there are various kinds of error sources, such as multipath fading, antenna mismatch, pulse distortion, etc. Plugging (6.55) and (6.56) into (6.5), the entries of the covariance matrix $\boldsymbol{\Sigma}$ are given by

$$[\boldsymbol{\Sigma}]_{i,j} = 4d_i d_j E[n_i n_j] + E[n_i^2 n_j^2] - \sigma_i^2 \sigma_j^2$$
$$= \begin{cases} 4d_i^2 \sigma_i^2 + 2\sigma_i^4 \approx \dfrac{12c^2}{N\pi^2\kappa^2} \dfrac{d_i^4}{\text{SNR}}, & i = j \\ 0, & i \neq j \end{cases}. \tag{6.57}$$

On the other hand, the TDOA estimates can be achieved by cross-correlating two received signals as follows

$$\hat{\tau}_{i,j} = \operatorname*{argmax}_{\tau_{i,j}} \left\{ \sum_{n=0}^{N-1} z_i(n) z_j(n - \tau_{i,j}) \right\}. \tag{6.58}$$

Thus, the estimate of the distance difference is $r_{i,j} = c\hat{\tau}_{i,j} = d_j - d_i + n_{i,j}$, where the bias is canceled out naturally. The statistical properties of $n_{i,j}$ can also be derived in a similar way as in [105, 132], and are given by

$$E[n_{i,j}] = 0, \qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.59)$$

$$\text{cov}(n_{i,j}, n_{p,q}) = \begin{cases} \dfrac{3c^2}{N\pi^2\kappa^2}\left(\dfrac{d_i^2}{\text{SNR}} + \dfrac{d_j^2}{\text{SNR}} + \dfrac{d_i^2 d_j^2}{\text{SNR}^2}\right) & i = p \text{ and } j = q \\[2ex] \dfrac{3c^2}{N\pi^2\kappa^2}\dfrac{d_i^2}{\text{SNR}} & i = p \text{ and } j \neq q \\[2ex] \dfrac{3c^2}{N\pi^2\kappa^2}\dfrac{d_j^2}{\text{SNR}} & j = q \text{ and } i \neq p \\[2ex] -\dfrac{3c^2}{N\pi^2\kappa^2}\dfrac{d_i^2}{\text{SNR}} & i = q \text{ and } j \neq p \\[2ex] -\dfrac{3c^2}{N\pi^2\kappa^2}\dfrac{d_j^2}{\text{SNR}} & j = p \text{ and } i \neq q \\[2ex] 0 & \text{else} \end{cases} .$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.60)$$

Note that similarly as in [105] the signal attenuation is taken into account in order to obtain more general noise statistics than in [132], but we correct the derivation errors in [105]. We remark that in reality, the TDOA estimates may face similar problems as the TOA estimates, since the received signals at different anchors may be totally different. Plugging (6.59) and (6.60) into (6.43), the entries of the covariance matrix $\boldsymbol{\Sigma}_i$ are given by

$$[\boldsymbol{\Sigma}_i]_{k,l} \approx \begin{cases} 4d_k d_l E[n_{i,k}n_{i,l}], & k < i \text{ and } l < i \\ 4d_{k+1}d_{l+1}E[n_{i,k+1}n_{i,l+1}], & k \geq i \text{ and } l \geq i \\ 4d_k d_{l+1}E[n_{i,k}n_{i,l+1}], & k < i \text{ and } l \geq i \\ 4d_{k+1}d_l E[n_{i,k+1}n_{i,l}], & k \geq i \text{ and } l < i \end{cases}$$

$$= \begin{cases} \dfrac{12c^2 d_k^2}{N\pi^2\kappa^2}\left(\dfrac{d_i^2}{\text{SNR}} + \dfrac{d_k^2}{\text{SNR}} + \dfrac{d_i^2 d_k^2}{\text{SNR}^2}\right), & k = l \text{ and } k < i \\[2ex] \dfrac{12c^2 d_{k+1}^2}{N\pi^2\kappa^2}\left(\dfrac{d_i^2}{\text{SNR}} + \dfrac{d_{k+1}^2}{\text{SNR}} + \dfrac{d_i^2 d_{k+1}^2}{\text{SNR}^2}\right), & k = l \text{ and } k \geq i \\[2ex] \dfrac{12c^2 d_k d_l}{N\pi^2\kappa^2}\dfrac{d_i^2}{\text{SNR}}, & k \neq l, \ k < i \text{ and } l < i \\[2ex] \dfrac{12c^2 d_{k+1}d_{l+1}}{N\pi^2\kappa^2}\dfrac{d_i^2}{\text{SNR}}, & k \neq l, \ k \geq i \text{ and } l \geq i \\[2ex] \dfrac{12c^2 d_k d_{l+1}}{N\pi^2\kappa^2}\dfrac{d_i^2}{\text{SNR}}, & k < i \text{ and } l \geq i \\[2ex] \dfrac{12c^2 d_{k+1}d_l}{N\pi^2\kappa^2}\dfrac{d_i^2}{\text{SNR}}, & k \geq i \text{ and } l < i \end{cases} .$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (6.61)$$

In the simulations, we generate n_i and $n_{i,j}$ as zero-mean Gaussian random vari-

ables with covariance matrices specified as above.

6.4.2   Performance evaluation

As a well-adopted lower bound, the Cramèr-Rao bound (CRB) is derived for localization estimators based on TOA measurements and TDOA measurements, respectively. Note that the estimators derived in this chapter are biased. We remark that although the CRB is a bound for unbiased estimators, it still is interesting to compare it with the proposed biased estimators. Here, we exemplify the CRBs for location estimation on a plane, e.g., we take $l = 2$. We assume that n_i and $n_{i,j}$ are Gaussian distributed. The Fisher information matrix (FIM) $\mathbf{I}_1(\boldsymbol{\theta})$ based on model (6.1) in Section 6.2 for TOA measurements is derived in Appendix 6.D, where $\boldsymbol{\theta} = [\mathbf{x}^T, \ b]^T$, and $\mathbf{x} = [x_1, \ x_2]^T$. Consequently, we obtain $\text{CRB}(x_1) = [\mathbf{I}_1^{-1}(\boldsymbol{\theta})]_{1,1}$. We observe that b is not part of $\mathbf{I}_1^{-1}(\boldsymbol{\theta})$. Therefore, no matter how large b is, it has the same influence on the CRB for TOA measurements. The FIM $\mathbf{I}_2(\mathbf{x})$ and $\mathbf{I}_3(\mathbf{x})$ based on model (6.39) in Section 6.3 are derived in Appendix 6.E for the nonredundant set and the full set of TDOA measurements, respectively.

We consider three simulation setups. In Setup 1 and Setup 2, eight anchors are evenly located on the edges of a $100\,\text{m} \times 100\,\text{m}$ rectangular. Meanwhile the target node is located at $[200\,\text{m}, 30\,\text{m}]$ and $[10\,\text{m}, 20\,\text{m}]$ for Setup 1 and Setup 2, respectively. Thus, the target node is far away from the anchors in Setup 1, but close to them in Setup 2. In Setup 3, all anchors and the target node are randomly distributed on a grid with cells of size $1\,\text{m} \times 1\,\text{m}$ inside the rectangular. The performance criterion is the root mean squared error (RMSE) of $\hat{\mathbf{x}}$ vs. a reference range SNR ($\text{SNR}_\text{r} = \frac{N\pi^2\kappa^2}{3c^2}\text{SNR}$), which can be expressed as $\sqrt{1/N_{exp} \sum_{j=1}^{N_{exp}} \|\hat{\mathbf{x}}^{(j)} - \mathbf{x}\|^2}$, where $\hat{\mathbf{x}}^{(j)}$ is the estimate obtained in the jth trial. Each simulation result is averaged over $N_{exp} = 1000$ Monte Carlo trials. The bias b corresponding to the clock offset is randomly generated in the range of $[0\,\text{m}, 100\,\text{m}]$ in each Monte Carlo run. We would like to compare all the REFF and REFB estimators, as well as the estimator proposed in [127] (first iteration) using TOA measurements, labeled the LS1 estimator, and the estimator proposed in [105] using the full TDOA set, namely the REFF-LS2 estimator.

Estimators using TOA measurements

Fig. 6.1 shows the localization performance of the REFF estimators using TOA measurements under the three considered setups. The CRB $\mathbf{I}_1^{-1}(\boldsymbol{\theta})$ (the dotted line with "×" markers) is used as a benchmark. The REFF-USR-WLS estimator (6.8) with the optimal weighting matrix (the solid line with "+" markers) achieves the best performance, while the iterative approach to update the weighting matrix (the solid line with "◇" markers) also helps the REFF-USR-WLS estimator to converge to the best performance. The REFF-SR-CLS estimator (6.14) (the solid line with "○" markers) benefits from the constraints, and thus outperforms the
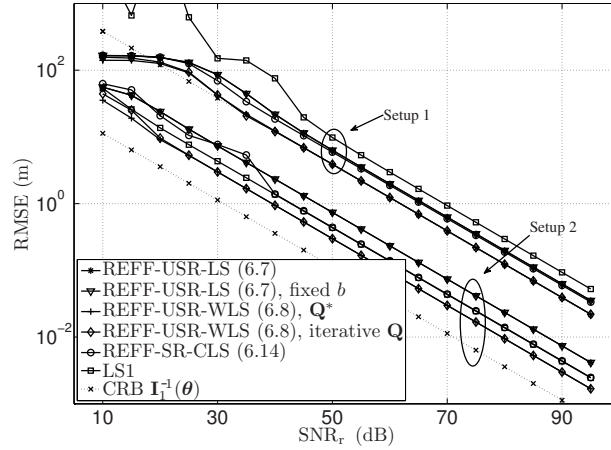
(a) Setup 1 and Setup 2
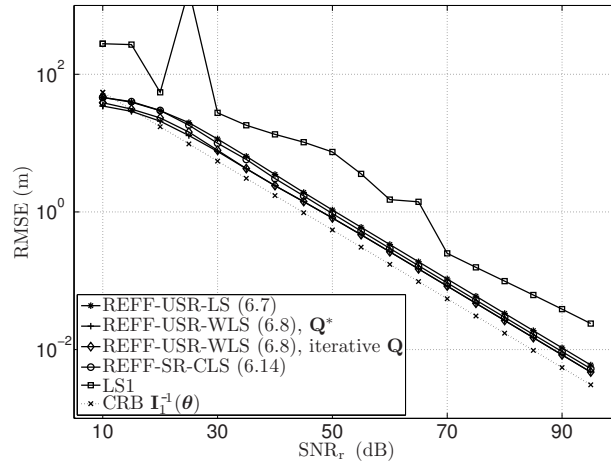
(b) Setup 3

Figure 6.1: RMSE of $\mathbf{x}$ for the REFF estimators using TOAs for locating an asynchronous target.

REFF-USR-LS estimator (6.7) (the solid line with "$*$" markers). The concrete value of the bias b does not influence the localization performance. The curve of the REFF-USR-LS estimator with fixed b (the solid line with "$\bigtriangledown$" markers) and the one with random b overlap. Furthermore, the LS1 estimator [127] (the solid line with "$\square$" markers) is sensitive to the geometry. It performs better than the REFF-USR-LS estimator in Setup 2, but worse in Setup 1. This observation is
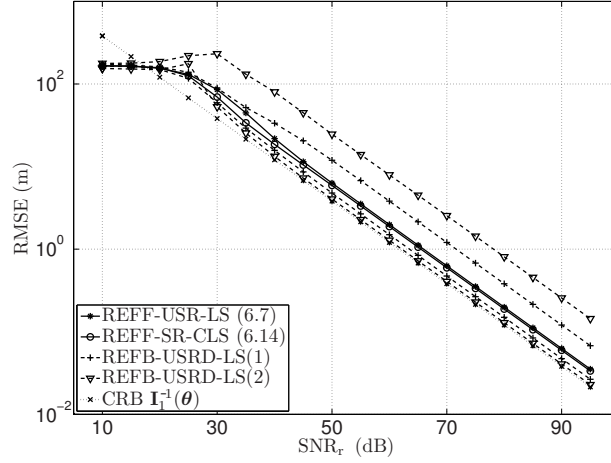
(a) Setup 1
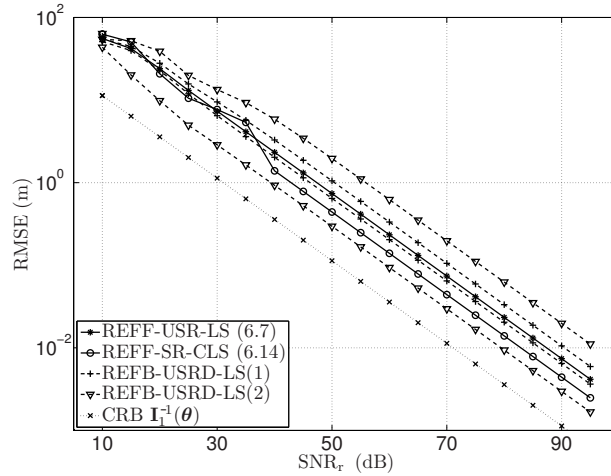
(b) Setup 2

Figure 6.2: RMSE of $\mathbf{x}$ for the REFF and the REFB estimators using TOAs for locating an asynchronous target.

consistent with the one in [120]. In Setup 3 (random geometry), it fails under some cases due to its inherent instability, and performs unsatisfactorily.

Fig. 6.2 compares the localization performance of the REFF with the one of the REFB estimators using TOA measurements under Setup 1 and Setup 2. Since there are no fixed anchors in Setup 3, we skip it in the comparison. We show both the performance of the best and the worst reference selection, which indicates the
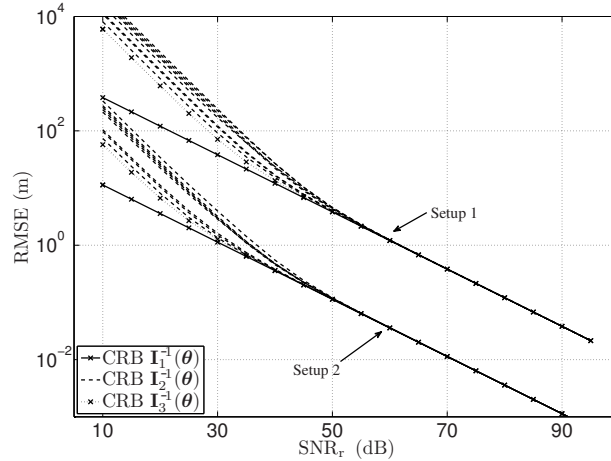
Figure 6.3: The CRBs using TOAs $\mathbf{I}_1^{-1}(\boldsymbol{\theta})$, the nonredundant TDOA set $\mathbf{I}_2^{-1}(\boldsymbol{\theta})$, and the full TDOA set $\mathbf{I}_3^{-1}(\boldsymbol{\theta})$ for locating an asynchronous target.

performance limits of the REFB estimators. The dashed lines with "+" and "$\bigtriangledown$" markers denote the performance bounds for the REFB-USRD-LS(1) and the REFB-USRD-LS(2), respectively. The best reference choice for the REFB-USRD-LS(1) estimator is the reference anchor with the shortest distance to the target node. Meanwhile, we do not observe the best reference pair selection for the REFB-USRD-LS(2) estimator following any rules. The curves for the REFF-USR-LS estimator (6.7) (the solid line with "*" markers) and the REFF-SR-CLS estimator (6.14) (the solid line with "∘" markers) lie inside these limits. Their performances are neither too bad nor too good, but they do not suffer from a poor reference selection. As we have already proved that the optimal weighting matrix can compensate the impact of the reference selection, the curves of all the WLS estimators with optimal weights will overlap. Thus, we do not show the performance of the REFF-USR-WLS estimator again, which is already illustrated in Fig. 6.1.

Estimators using TDOA measurements

Let us first compare the CRBs employing different measurements in Fig. 6.3. We observe the same tendency for both Setup 1 and Setup 2. All the CRBs overlap above a specific $\mathrm{SNR_r}$ threshold, which is 55 dB for Setup 1, and 50 dB for Setup 2. Below the threshold, the CRB using TOA measurements (the solid line with "×" markers) is lower than the other CRBs. Meanwhile, the CRB using the full TDOA set (the dotted line with "×" markers) is lower than the ones using a nonredundant TDOA set (the dotted lines). The observations are consistent with the ones in [132]. On the other hand, the $\mathrm{SNR_r}$ ranges of interest corresponding
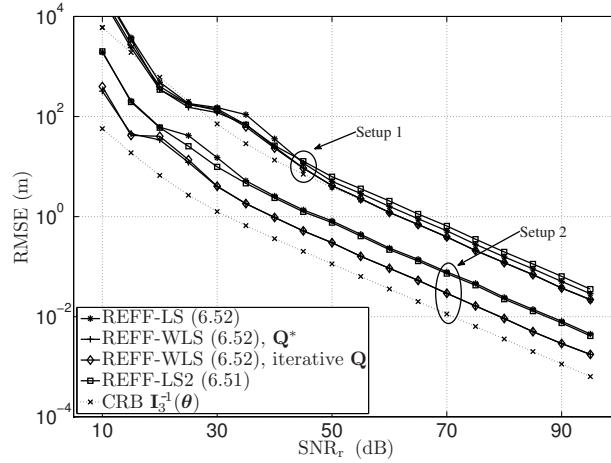
(a) Setup 1 and Setup 2
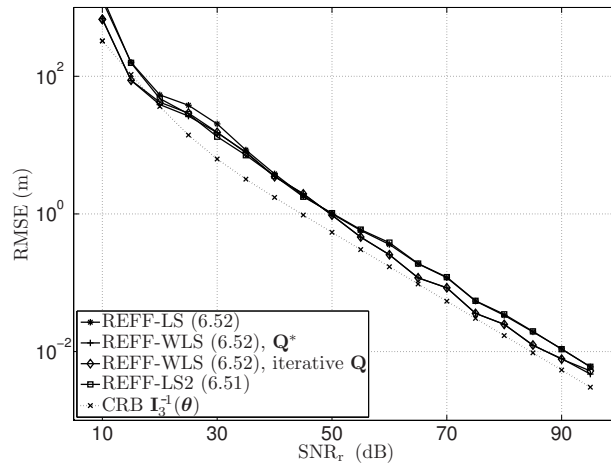
(b) Setup 3

Figure 6.4: RMSE of $\mathbf{x}$ for the REFF estimators using the full set of TDOAs for locating an asynchronous target.

to a RMSE smaller than $10^0 = 1\,\text{m}$, are $\text{SNR}_r > 60\,\text{dB}$ and $\text{SNR}_r > 30\,\text{dB}$ for Setup 1 and Setup 2, respectively. Within this range of interest, there are no differences among the CRBs in Setup 1, and only small differences in Setup 2. Therefore, using different measurements would not cause obvious differences in the CRB at high SNR.

Fig. 6.4 shows the localization performance of the REFF estimators using

the full TDOA set under three setups. The CRB $\mathbf{I}_3^{-1}(\boldsymbol{\theta})$ (the dotted line with "$\times$" markers) is still used as a benchmark. We observe similar tendencies as in Fig. 6.1. The REFF-WLS estimator based on (6.52) with the optimal weighting matrix (the solid line with "$+$" markers) achieves the best performance, while the iterative approach to update the weighting matrix (the solid line with "$\diamondsuit$" markers) also facilitates the REFF-WLS estimator based on (6.52) to converge to the best performance. Moreover, the performance of the REFF-LS2 estimator (6.51) [105] (the solid line with "$\square$" markers) is slightly worse than the REFF-LS estimator based on (6.52) (the solid line with "$*$" markers) in Setup 1. In general, their performances are very close. In Setup 3 (random geometry), they almost overlap with each other.

Fig. 6.5 compares the localization performance of the REFF estimator using the full TDOA set with the one of the REFB estimators using the nonredundant TDOA set under Setup 1 and Setup 2. Since there are no fixed anchors in Setup 3, we again skip it in the comparison. We show both the performance of the best and the worst reference selection, which indicates the performance limits of the REFB estimators. The dashed lines with "$+$" and "$\triangledown$" markers denote the performance limits for the REFB-USRD-LS(1) (6.44) and the REFB-USRD-LS(2) estimator, respectively. The best reference choice for the REFB-USRD-LS(1) estimator is again the reference anchor with the shortest distance to the target node, which means we cross-correlate the received signal at the reference anchor with the ones at other anchors in order to achieve a nonredundant set of TDOA measurements. Meanwhile, we do not observe the best reference pair selection for the REFB-USRD-LS(2) estimator following any rules either. The curves for the REFF-LS estimator based on (6.52) (the solid line with "$*$" markers) and the REFF-LS2 estimator (6.51) [105] (the solid line with "$\square$" markers) lie inside these limits. They are very close to the lower limits in Setup 1, and in the middle of the performance band in Setup 2. The performance band of the REFB-USRD-LS(1) estimator is quite narrow in Setup 2. On the other hand, the performance variation is very obvious for the REFB-USRD-LS(2) estimator.

Finally, we verify the equivalence of the REFB-USRD-WLS estimators with the same optimal weighting matrix in Fig. 6.6. As we have discussed before, the optimal weighting matrix can only release the impact of the second reference selection. The first reference selection decides the obtained data set. Therefore, using the same nonredundant set of TDOAs, the curves of the REFB-USRD-WLS(1) (6.45) (the solid lines with "$\diamondsuit$" markers) and the REFB-USRD-WLS(2) estimators (the solid lines with "$+$" markers) overlap. A different performance can be obtained by employing different nonredundant TDOA sets. However, similarly as the CRB, the performance converges after some $\mathrm{SNR_r}$ threshold. Finally, in Fig. 6.7, we compare the localization performance of the REFF estimators using TOAs and the full TDOA set, respectively. They are very close at high $\mathrm{SNR_r}$, but diverge at low $\mathrm{SNR_r}$.
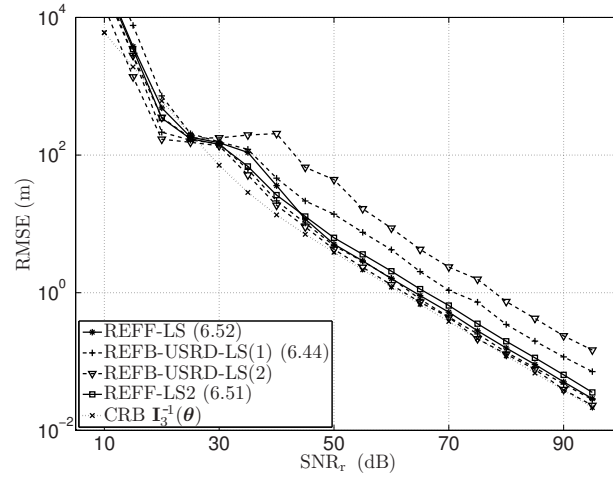
(a) Setup 1
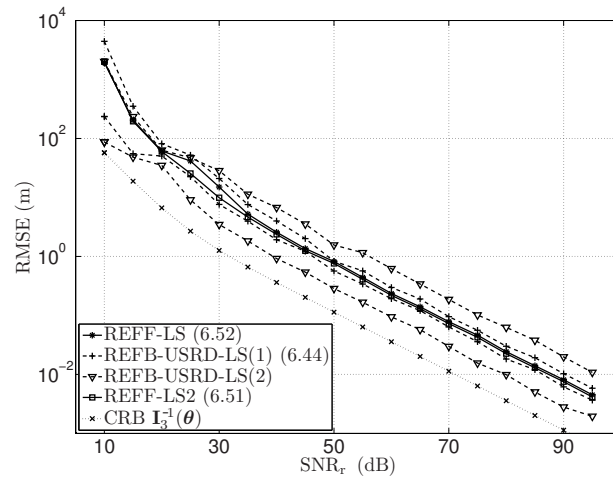
(b) Setup 2

Figure 6.5: RMSE of $\mathbf{x}$ for the REFF estimator using the full set of TDOAs and the REFB estimators using the nonredundant set of TDOAs for locating an asynchronous target.

6.5 Conclusions

In this chapter, we have proposed reference-free localization estimators based on TOA measurements for a scenario, where anchors are synchronized, and the clock of the target node runs freely. The reference-free estimators do not suffer from a
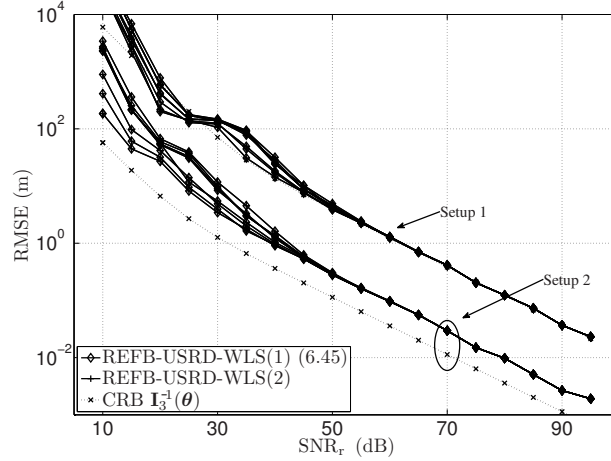
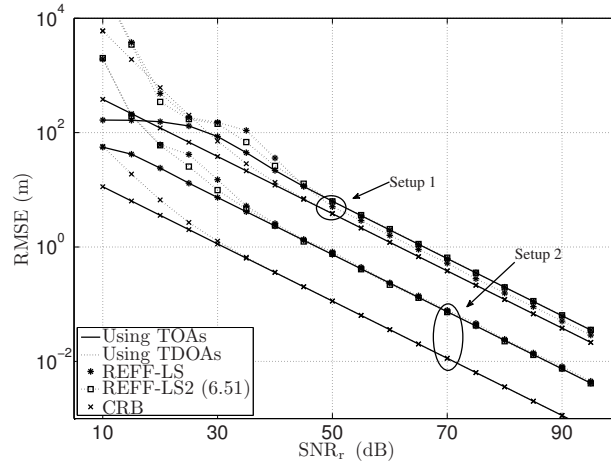**Figure 6.6:** RMSE of $\mathbf{x}$ for the REFB-USRD-WLS estimators using the nonredundant set of TDOAs for locating an asynchronous target.

Figure 6.7: RMSE of $\mathbf{x}$ for the REFF estimator using TOAs and the full set of TDOAs for locating an asynchronous target.

poor reference selection, which can seriously degrade the localization performance of reference-based LS estimators. Furthermore, we generalized existing reference-based localization estimators using TOA or TDOA measurements, and expose their relations. Based on analysis and simulations, we have obtained the following important conclusions:

1. Applying a projection is always preferred over making differences with a

reference to get rid of nuisance parameters.

2. The optimal weighting matrix can compensate for the impact of the reference selection for reference-based WLS estimators using TOA measurements, and make all those estimators equivalent. However, the optimal weighting matrix cannot release the reference influence for reference-based WLS estimators using a nonredundant set of TDOA measurements, but can make the estimators using the same set identical as well.

3. There are corresponding equivalences between the squared-range-based and the squared-range-difference-based methods, which are all using TOA measurements.

4. Beyond some SNR threshold, there are no obvious differences among the CRBs using TOA measurements, the nonredundant set and the full set of TDOA measurements, respectively.

5. The performance of the reference-free LS estimators is neither too bad nor too good, but they do not suffer from a poor reference selection.

6. The concrete value of the distance bias caused by the inaccurate clock does not affect the localization performance of the LS or WLS estimators.

6.A   Derivation of λ for CLS

Substituting (6.14) into the constraint (6.11), we obtain

$$(\boldsymbol{\phi}^T\mathbf{W}^T\mathbf{A} - \frac{\lambda}{2}\boldsymbol{\rho}^T)(\mathbf{A}^T\mathbf{W}\mathbf{A} + \lambda\mathbf{J})^{-1}\mathbf{J}(\mathbf{A}^T\mathbf{W}\mathbf{A} + \lambda\mathbf{J})^{-1}(\mathbf{A}^T\mathbf{W}\boldsymbol{\phi} - \frac{\lambda}{2}\boldsymbol{\rho})$$
$$+\boldsymbol{\rho}^T(\mathbf{A}^T\mathbf{W}\mathbf{A} + \lambda\mathbf{J})^{-1}(\mathbf{A}^T\mathbf{W}\boldsymbol{\phi} - \frac{\lambda}{2}\boldsymbol{\rho}) = 0, \quad (6.62)$$

which has to be solved for λ, leading to the estimate $\hat{\lambda}$. We exemplify how to solve (6.62) for localization on a plane, i.e., $l = 2$. Since $\mathbf{J}$ is of rank 3, there are only three non-zero eigenvalues of $(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{J}$. Therefore, the square matrix $(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{J}$ of size 4×4 can be diagonalized as $(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{J} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}^{-1}$, where $\mathbf{V}$ is of size 4×3, collecting the singular vectors corresponding to the three nonzero singular values, and $\boldsymbol{\Lambda}$ is a diagonal matrix with the three nonzero singular values $(\gamma_i, i = 1, 2, 3)$ on its diagonal. According to the Kailath variant [104] and plugging the eigenvalue decomposition of $(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{J}$ into $(\mathbf{A}^T\mathbf{W}\mathbf{A} + \lambda\mathbf{J})^{-1}$, we obtain

$$
\begin{aligned}
&(\mathbf{A}^T\mathbf{W}\mathbf{A} + \lambda\mathbf{J})^{-1} \\
=\ & (\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1} - \lambda(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{J}(\mathbf{I} + \lambda(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{J})^{-1}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1} \\
=\ & \mathbf{V}(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{V}^{-1}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1} \quad\quad\quad (6.63)
\end{aligned}
$$

Substituting (6.63) into the constraint (6.62), we achieve

$$
\begin{aligned}
0\ =\ & \mathbf{e}^T(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{f} + \frac{\lambda^2}{4}\mathbf{h}^T(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{g} \\
& -\frac{\lambda}{2}\mathbf{h}^T(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{f} - \frac{\lambda}{2}\mathbf{e}^T(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\boldsymbol{\Lambda}(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{g} \\
& +\mathbf{h}^T(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{f} - \frac{\lambda}{2}\mathbf{h}^T(\mathbf{I} + \lambda\boldsymbol{\Lambda})^{-1}\mathbf{g} \quad\quad\quad (6.64)
\end{aligned}
$$

where

$$
\begin{aligned}
\mathbf{e}^T\ &=\ \boldsymbol{\phi}^T\mathbf{W}^T\mathbf{A}\mathbf{V} = [e_1 \quad e_2 \quad e_3 \quad e_4], & (6.65) \\
\mathbf{f}\ &=\ \mathbf{V}^{-1}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{W}\boldsymbol{\phi} = [f_1 \quad f_2 \quad f_3 \quad f_4]^T, & (6.66) \\
\mathbf{h}^T\ &=\ \boldsymbol{\rho}^T\mathbf{V} = [h_1 \quad h_2 \quad h_3 \quad h_4], & (6.67) \\
\mathbf{g}\ &=\ \mathbf{V}^{-1}(\mathbf{A}^T\mathbf{W}\mathbf{A})^{-1}\boldsymbol{\rho} = [g_1 \quad g_2 \quad g_3 \quad g_4]^T. & (6.68)
\end{aligned}
$$

Now, (6.64) can be simplified as a seven-order equation as follows

$$
\begin{aligned}
0 &= \sum_{i=1}^{3} \frac{e_i f_i \gamma_i}{(1+\lambda\gamma_i)^2} + \frac{\lambda^2}{4}\sum_{i=1}^{3}\frac{h_i g_i \gamma_i}{(1+\lambda\gamma_i)^2} - \frac{\lambda}{2}\sum_{i=1}^{3}\frac{e_i g_i \gamma_i}{(1+\lambda\gamma_i)^2} - \frac{\lambda}{2}\sum_{i=1}^{3}\frac{h_i f_i \gamma_i}{(1+\lambda\gamma_i)^2} \\
&\quad + \sum_{i=1}^{3}\frac{h_i f_i \gamma_i}{(1+\lambda\gamma_i)} - \frac{\lambda}{2}\sum_{i=1}^{3}\frac{h_i g_i \gamma_i}{(1+\lambda\gamma_i)} + h_4 f_4 - \frac{\lambda}{2}h_4 g_4.
\end{aligned}
\tag{6.69}
$$

After obtaining the seven roots of (6.69), we discard the complex roots, and plug the real roots into (6.14). Finally, we choose the estimate $\hat{\mathbf{y}}$, which fulfills (6.10). Note that (6.14) is a CLS estimate of $\mathbf{y}$ with $\mathbf{W} = \mathbf{I}$. Since the optimal $\mathbf{W}^*$ depends on the unknown $\mathbf{d}$, the CWLS problem can be solved in a similar way by iteratively updating the weights and the estimates, thus we do not repeat it here.

6.B   Proof of $\mathbf{P}_i((\mathbf{T}_i\mathbf{u}) \odot (\mathbf{T}_i\mathbf{u})) = \mathbf{P}_i\mathbf{T}_i(\mathbf{u} \odot \mathbf{u})$

Recalling that $\mathbf{T}_i\mathbf{u} = \mathbf{T}_{i1}\mathbf{u} - u_i\mathbf{1}_{M-1}$, $\mathbf{T}_i\mathbf{1}_M = \mathbf{0}_{M-1}$, and $\mathbf{P}_i\mathbf{T}_i\mathbf{u} = \mathbf{0}_{M-1}$, we prove that $\mathbf{P}_i((\mathbf{T}_i\mathbf{u}) \odot (\mathbf{T}_i\mathbf{u}))$ in (6.33) is equivalent to $\mathbf{P}_i\mathbf{T}_i(\mathbf{u} \odot \mathbf{u})$ as follows

$$
\begin{aligned}
\mathbf{P}_i((\mathbf{T}_i\mathbf{u}) \odot (\mathbf{T}_i\mathbf{u})) &= \mathbf{P}_i((\mathbf{T}_{i1}\mathbf{u} - u_i\mathbf{1}_{M-1}) \odot (\mathbf{T}_{i1}\mathbf{u} - u_i\mathbf{1}_{M-1})) \\
&= \mathbf{P}_i((\mathbf{T}_{i1}\mathbf{u}) \odot (\mathbf{T}_{i1}\mathbf{u}) - 2u_i\mathbf{T}_{i1}\mathbf{u} + u_i^2\mathbf{1}_{M-1}) \\
&= \mathbf{P}_i((\mathbf{T}_{i1}\mathbf{u}) \odot (\mathbf{T}_{i1}\mathbf{u}) - 2u_i(\mathbf{T}_{i1}\mathbf{u} - u_i\mathbf{1}_{M-1}) - u_i^2\mathbf{1}_{M-1}) \\
&= \mathbf{P}_i((\mathbf{T}_{i1}\mathbf{u}) \odot (\mathbf{T}_{i1}\mathbf{u}) - 2u_i\mathbf{T}_i\mathbf{u} - u_i^2\mathbf{1}_{M-1}) \\
&= \mathbf{P}_i((\mathbf{T}_{i1}\mathbf{u}) \odot (\mathbf{T}_{i1}\mathbf{u}) - u_i^2\mathbf{1}_{M-1}) \\
&= \mathbf{P}_i\mathbf{T}_i(\mathbf{u} \odot \mathbf{u})
\end{aligned}
$$

6.C   Derivation of (6.35)

The singular value decomposition (SVD) of $\mathbf{P}_i\mathbf{T}_i$ is given by $\mathbf{P}_i\mathbf{T}_i = \mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^T$, where $\mathbf{U}_i$ is of size $(M-1)\times(M-2)$ and $\mathbf{V}_i$ is of size $M\times(M-2)$, which collect the left and right singular vectors corresponding to the $M-2$ nonzero singular values, and $\mathbf{\Lambda}_i$ is a diagonal matrix with the $M-2$ nonzero singular values on its diagonal. Note that $\mathbf{U}_i^T\mathbf{U}_i = \mathbf{I}_{M-2}$, $\mathbf{V}_i^T\mathbf{V}_i = \mathbf{I}_{M-2}$, $\mathbf{V}_i^T\mathbf{1}_M = \mathbf{0}_{M-2}$ and $\mathbf{V}_i^T\mathbf{u} = \mathbf{0}_{M-2}$. As a result, the nullspace $\mathcal{N}(\mathbf{V}_i^T) = \mathrm{span}(\mathbf{1}_M, \mathbf{u})$, and $\mathcal{R}(\mathbf{V}_i) = \mathcal{R}(\mathbf{U})$. Using the SVD and the property of the pseudo-inverse, we can write $(\mathbf{P}_i\mathbf{T}_i\mathbf{\Sigma}\mathbf{T}_i^T\mathbf{P}_i)^\dagger$ as

$$
\begin{aligned}
(\mathbf{P}_i\mathbf{T}_i\mathbf{\Sigma}\mathbf{T}_i^T\mathbf{P}_i)^\dagger &= (\mathbf{U}_i\mathbf{\Lambda}_i\mathbf{V}_i^T\mathbf{\Sigma}\mathbf{V}_i\mathbf{\Lambda}_i\mathbf{U}_i^T)^\dagger \\
&= (\mathbf{\Lambda}_i\mathbf{U}_i^T)^\dagger(\mathbf{V}_i^T\mathbf{\Sigma}\mathbf{V}_i)^{-1}(\mathbf{U}_i\mathbf{\Lambda}_i)^\dagger \\
&= \mathbf{U}_i\mathbf{\Lambda}_i^{-1}(\mathbf{V}_i^T\mathbf{\Sigma}\mathbf{V}_i)^{-1}\mathbf{\Lambda}_i^{-1}\mathbf{U}_i^T.
\end{aligned}
\tag{6.70}
$$

Plugging (6.70) and the SVD of $\mathbf{P}_i\mathbf{T}_i$ into (6.34), and making use of the property of the pseudo-inverse again, we arrive at

$$
\begin{aligned}
\mathbf{T}_i^T\mathbf{P}_i(\mathbf{P}_i\mathbf{T}_i\boldsymbol{\Sigma}\mathbf{T}_i^T\mathbf{P}_i)^\dagger\mathbf{P}_i\mathbf{T}_i &= \mathbf{V}_i\boldsymbol{\Lambda}_i\mathbf{U}_i^T\mathbf{U}_i\boldsymbol{\Lambda}_i^{-1}(\mathbf{V}_i^T\boldsymbol{\Sigma}\mathbf{V}_i)^{-1}\boldsymbol{\Lambda}_i^{-1}\mathbf{U}_i^T\mathbf{U}_i\boldsymbol{\Lambda}_i\mathbf{V}_i^T \\
&= \mathbf{V}_i(\mathbf{V}_i^T\boldsymbol{\Sigma}\mathbf{V}_i)^{-1}\mathbf{V}_i^T \\
&= (\mathbf{V}_i\mathbf{V}_i^T\boldsymbol{\Sigma}\mathbf{V}_i\mathbf{V}_i^T)^\dagger,
\end{aligned}
\tag{6.71}
$$

where $\mathbf{V}_i\mathbf{V}_i^T$ is the projection onto $\mathcal{R}(\mathbf{U})$.

6.D   CRB derivation for localization based on TOA measurements

We analyze the CRB for jointly estimating $\mathbf{x}$ and b based on (6.1), and assume n_i is Gaussian distributed. The FIM $\mathbf{I}_1(\boldsymbol{\theta})$ is employed, where $\boldsymbol{\theta} = [\mathbf{x}^T,\ b]^T$, with entries defined as:

$$
\begin{aligned}
\mathbf{I}_1(\boldsymbol{\theta}) &= -E\left[\frac{\partial^2\ln p(\mathbf{u};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^T}\right] \\
&= \left[\frac{\partial\boldsymbol{\nu}}{\partial\boldsymbol{\theta}}\right]^T\mathbf{C}^{-1}\left[\frac{\partial\boldsymbol{\nu}}{\partial\boldsymbol{\theta}}\right],
\end{aligned}
\tag{6.72}
$$

where

$$
\begin{aligned}
\boldsymbol{\nu} &= \mathbf{d} + b\mathbf{1}_M, \tag{6.73} \\
\mathbf{C} &= \mathrm{diag}([\sigma_1^2,\ \sigma_2^2,\ \ldots,\ \sigma_M^2]^T), \\
&= \frac{3c^2}{N\pi^2\kappa^2\mathrm{SNR}}\mathrm{diag}([d_1^2,\ d_2^2,\ \ldots,\ d_M^2]^T), \tag{6.74} \\
\frac{\partial\boldsymbol{\nu}}{\partial b} &= \mathbf{1}_M, \tag{6.75} \\
\left[\frac{\partial\boldsymbol{\nu}}{\partial x_l}\right]_j &= \frac{x_l - x_{l,j}}{\|\mathbf{x} - \mathbf{x}_j\|}. \tag{6.76}
\end{aligned}
$$

6.E   CRB derivation for localization based on TDOA measurements

We analyze the CRB for estimating $\mathbf{x}$ based on (6.40), and assume $n_{i,j}$ is Gaussian distributed. The FIM $\mathbf{I}_2(\mathbf{x})$ for the nonredundant set of TDOA measurements is

employed, with entries defined as:

$$
\begin{aligned}
\mathbf{I}_2(\mathbf{x}) &= -E\left[\frac{\partial^2 \mathrm{ln}p(\mathbf{r}_i;\mathbf{x})}{\partial \mathbf{x} \partial \mathbf{x}^T}\right] \\
&= \left[\frac{\partial \boldsymbol{\mu}_i}{\partial \mathbf{x}}\right]^T \mathbf{C}_i^{-1} \left[\frac{\partial \boldsymbol{\mu}_i}{\partial \mathbf{x}}\right],
\end{aligned}
\tag{6.77}
$$

where

$$
\boldsymbol{\mu}_i = \mathbf{T}_{i1}\mathbf{d} - d_i \mathbf{1}_{M-1},
\tag{6.78}
$$

$$
[\mathbf{C}_i]_{k,l} =
\begin{cases}
\dfrac{3c^2}{N\pi^2\kappa^2}\left(\dfrac{d_i^2}{\mathrm{SNR}} + \dfrac{d_k^2}{\mathrm{SNR}} + \dfrac{d_i^2 d_k^2}{\mathrm{SNR}^2}\right) & k = l \text{ and } k < i \\[3mm]
\dfrac{3c^2}{N\pi^2\kappa^2}\left(\dfrac{d_i^2}{\mathrm{SNR}} + \dfrac{d_{k+1}^2}{\mathrm{SNR}} + \dfrac{d_i^2 d_{k+1}^2}{\mathrm{SNR}^2}\right) & k = l \text{ and } k \geq i \\[3mm]
\dfrac{3c^2}{N\pi^2\kappa^2}\dfrac{d_i^2}{\mathrm{SNR}} & \text{else}
\end{cases}
\tag{6.79}
$$

$$
\left[\frac{\partial \boldsymbol{\mu}_i}{\partial x_j}\right]_k =
\begin{cases}
\dfrac{x_j - x_{j,k}}{\|\mathbf{x} - \mathbf{x}_k\|} - \dfrac{x_j - x_{j,i}}{\|\mathbf{x} - \mathbf{x}_i\|}, & k < i \\[3mm]
\dfrac{x_j - x_{j,k+1}}{\|\mathbf{x} - \mathbf{x}_{k+1}\|} - \dfrac{x_j - x_{j,i}}{\|\mathbf{x} - \mathbf{x}_i\|}, & k \geq i
\end{cases}.
\tag{6.80}
$$

Furthermore, let us define $\boldsymbol{\mu} = [\boldsymbol{\mu}_1^T, [\boldsymbol{\mu}_2^T]_{2:M}, \ldots, [\boldsymbol{\mu}_{M-1}^T]_{M-1}]^T$, where $\boldsymbol{\mu}_i = [\mu_{i,1}, \ldots, \mu_{i,i-1}, \mu_{i,i+1}, \ldots, \mu_{i,M}]^T$, and $\mathbf{C}$ as the covariance matrix of this full set of TDOA measurements. Then the FIM $\mathbf{I}_3(\mathbf{x})$ for the full set can also be derived based on (6.77) by replacing $\boldsymbol{\mu}_i$ and $\mathbf{C}_i$ with $\boldsymbol{\mu}$ and $\mathbf{C}$, respectively. We can obtain $[\boldsymbol{\mu}]_k = \mu_{i,j}$, where $k = (i-1)M - i^2/2 - i/2 + j$, $k \in \{1, 2, \ldots, M(M-1)/2\}$, $i \in \{1, 2, \ldots, M-1\}$, $j \in \{2, 3, \ldots, M\}$ and $j > i$. Consequently, we achieve

$$
\left[\frac{\partial \boldsymbol{\mu}}{\partial x_l}\right]_k = \frac{x_l - x_{l,j}}{\|\mathbf{x} - \mathbf{x}_j\|} - \frac{x_l - x_{l,i}}{\|\mathbf{x} - \mathbf{x}_i\|}.
\tag{6.81}
$$

In the same way, $[\mathbf{C}]_{k,l} = \mathrm{cov}(n_{i,j}, n_{p,q})$, where $l = (p-1)M - p^2/2 - p/2 + q$, $l \in \{1, 2, \ldots, M(M-1)/2\}$, $p \in \{1, 2, \ldots, M-1\}$, $q \in \{2, 3, \ldots, M\}$ and $q > p$.

Robust Time-Based Localization

7.1 Introduction

In this chapter, we employ TOA measurements for our localization methods because of their high accuracy and potentially low cost implementation based on ultra-wideband impulse radios (UWB-IRs). Since TOA measurements are time-based, clock synchronization is essential to achieve accurate localization. Clock synchronization alone plays a critical role to guarantee general operations of WSNs. It is under intensive investigation [29, 139–141] and results in various protocols, such as the Reference Broadcast Synchronization (RBS) protocol [142], the Timing-sync Protocol for Sensor Networks (TPSN) [27], and the Flooding Time Synchronization Protocol (FTSP) [143]. On the other hand, clock synchronization can also be handled by signal processing tools: a maximum likelihood estimator (MLE) for the clock offset is designed in [144], whereas theoretical performance limits for clock synchronization and MLEs for the clock offset and skew under different delay models, are developed in [145] and [146].

Because of the stringent cost and power constraints of WSNs, low-cost clocks are normally employed. This makes time-based localization and synchronization tightly coupled and challenging [141]. However, only recently the two entangled problems are jointly considered. A time-based positioning scheme (TPS) is developed in [147], where only the clock offset is considered. In [148], location in time and space is proposed, but only at the MAC and application layers. The two-way ranging (TWR) protocol proposed in the IEEE 802.15.4a standard [75] is employed in [149] for asynchronous networks. The relative clock skews are first

The results in this chapter appeared in [135–138].

calibrated, and then the node positions are estimated by a distributed maximum log-likelihood estimator (MLLE). Furthermore, a localization approach based on triple-differences, which are the differences of two differential TDOAs, is proposed in [150], where the corrupted one-way TOA measurements due to the relative clock offset and clock skew are corrected by several steps. The TWR protocol is also employed in [151], which considers TOA-based localization using practical clocks with internal delays and clock skews. Only recently, the joint estimation of the clock skew, the clock offset, and the position of the target node is proposed in [28] for networks with synchronous anchors. A total least-squares (TLS) estimator is further proposed in [28] to take the uncertainties of the anchor positions and clock parameters into account. Moreover, the target node position and clock offset are estimated by a weighted least-squares (WLS) estimator in [127]. An asynchronous position measurement system is developed in [152] for indoor localization.

Due to the burgeoning of WSNs, localization is vulnerable to many types of attacks (see e.g., [22–24]). Lazos and Poovendran [153] propose Secure Range-independent Localization (SeRLoc) methods by taking advantage of antenna sectors in the presence of malicious adversaries. Capkun et al. [154] design Secure Positioning In sensor NEtworks (SPINE), which deal with distance modification attacks. Moreover, Chen et al. [155] design several attack detection schemes for wireless localization systems. Li et al. [118] propose to use least median squares (LMS) as the metric to develop localization algorithms, which tolerate outliers. Liu et al. [156] use the minimum mean square error (MMSE) as an indicator to filter out outliers, and further propose another method to bear with outliers by adopting an iteratively refined voting scheme.

In this chapter, we consider time-related attacks which tightly connect with localization and synchronization. UWB-IRs are employed for high resolution TOA ranging [1, 76]. Low duty cycle, low probability of detection and speed of light transmission make UWB-IRs ideal for secure communication and localization. The TWR protocol in the IEEE 802.15.4a standard [75] promotes UWB ranging. However, this TWR protocol is vulnerable to an internal ranging attack by deceitful target nodes, which means that target nodes can send fraudulent timestamps to spoof their processing time. Furthermore, target nodes may submit inaccurate timestamps due to their asynchronous clocks or other reasons. Thus, the current protocol is not efficient and can even fail under the above circumstances.

We adopt UWB transmissions and propose robust TOA-based localization methods for asynchronous networks with possible internal attacks. We deal with two kinds of networks w.r.t. the synchronization level of anchors: one with only clock offsets among the anchors, referred to as quasi-synchronous networks, and the other with not only clock offsets but also clock skews among the anchors, referred to as fully asynchronous networks. Note that there is no synchronization requirement for the target node, whose clock can run freely. A novel ranging protocol, namely the asymmetric trip ranging (ATR) protocol is proposed in this chapter by taking advantage of the broadcast property of WSNs. All the anchors can obtain ranging information in one ranging procedure. The ATR protocol

reduces the communication load dramatically compared to the TWR protocol. In addition, by ignoring the processing time report from the target node, we estimate the target node position based only on the reliable timestamps from the anchors. As a result, the fact that the target node is not synchronized to the anchors, or an internal attack is mounted by a compromised target node, does not have any influence on the performance of our method. Furthermore, closed-form least-squares (LS) (and WLS) estimators using the ATR protocol are proposed for quasi-synchronous and fully asynchronous networks, respectively. In the case of fully asynchronous networks using the ATR protocol, we develop two different methods. One aims to achieve a small communication load, and practical issues are taken into account to simplify the estimators. To further simplify these estimators, synchronization and localization can be decoupled with almost no cost in performance. The other method relaxes the communication load constraint in order to cancel the error floor at high SNR due to approximations in the first method. But it still has a smaller communication load compared to the TWR protocol. To the best of our knowledge, this is the first study that combines all three aspects: localization, synchronization and security. Moreover, we also propose closed-form LS estimators for fully asynchronous networks using the TWR protocol, which overcome some drawbacks of previous work in literatures.

The rest of the chapter is organized as follows. In Section 7.2, we review the TWR protocol, analyze its error sources and explain its vulnerability. The robust time-based localization algorithms for quasi-synchronous and fully asynchronous networks are proposed in Section 7.3 and Section 7.4, respectively. In Section 7.5, we propose a joint synchronization and localization method using the TWR protocol for fully asynchronous networks. Performance bounds and simulation results are shown in Section 7.6. The conclusions are drawn at the end of this chapter.
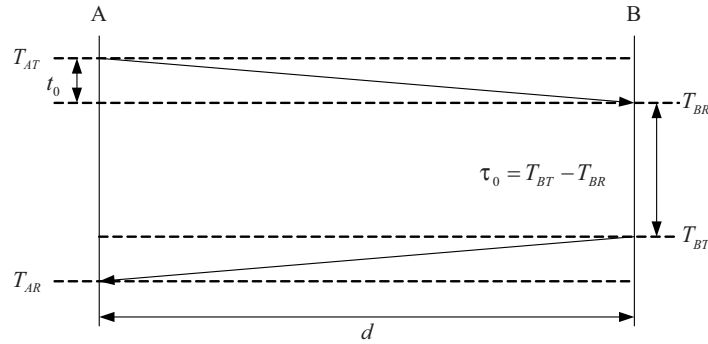
7.2   The TWR protocol

The TWR protocol used in the IEEE 802.15.4a standard [75] facilitates ranging between two nodes. The packet structure proposed by the standard is composed of a synchronization header (SHR) preamble, a physical layer header (PHR) and a data field (see Fig. 7.3). The first pulse of the PHR is called the ranging marker (RMARKER). The moment when the RMARKER leaves or arrives at the antenna of a node is critical to ranging. An example of the TWR protocol is shown in Fig. 7.1. Node A (or Node B) records T_{AT} (or T_{BT}) and T_{AR} (or T_{BR}) upon the departure and the arrival of the RMARKER, respectively. Thus, the time of flight (TOF) t_0, which is linear to the distance d (the ranging target) between node A and node B ($d = ct_0$, where c is the speed of light), is given by

$$t_0 \quad = \quad \frac{1}{2} \left(\frac{T_{AR} - T_{AT}}{\alpha_A} - \frac{\tau_0}{\alpha_B} \right) + n, \qquad\qquad (7.1)$$

where $\tau_0 = T_{BT} - T_{BR}$ is the processing time at node B, α_A and α_B are the clock skews of node A and node B, respectively, and n is the aggregate error term. In general, τ_0 is several hundreds of milliseconds and t_0 is several tens of nanoseconds for an indoor environment. As a result, τ_0 is heavily influenced by clock drift due to the relative clock skew between node A and node B.

As the differences of the timestamps are employed in (7.1), the clock offsets are canceled. The aggregate error term n in (7.1) accounts for different kinds of errors, such as the leading edge detection (LED) errors [75] [157], and the uncertainties of the internal delays [75, 151], which we explain next. The LED errors are due to the detection of the first multipath component of the received RMARKER. It is not a trivial task to detect the first multipath component, or in other words, the leading edge of the first cluster of the arriving RMARKER. It depends on the multipath channel, the signal bandwidth, the signal-to-noise ratio (SNR) and the detection strategy [1] (see our investigations in Part II). Even if we can detect the leading edge accurately, there may still be a ranging bias due to different kinds of environments. The LED could lead to accurate ranging for line-of-sight (LOS). However, the TOF of the first arriving multipath component in non-line-of-sight (NLOS) environments may not indicate the correct distance information due to the obstacles between two ranging nodes. In that case, the TOF estimates are biased, and calibration is indispensable to remove the bias. A ranging model that can distinguish between different environments is proposed in [157]. A more sophisticated ranging model as an enhancement of [157] is presented in [149], which takes the detection noise and the drift compensation into account. Next to LED errors, internal delays are caused by the difficulty to measure events at the antenna exactly. Since the ranging counter is typically somewhere in the digital section, the signal has to go through some transmitting (or receiving) chain after (or before) the ranging counter records the timestamp to reach the antenna (or the ranging counter). There is a difference between the real time the RMARKER leaves or arrives at the antenna and the recorded time by the ranging counter. This time-varying internal delay can be a few hundreds of nanoseconds depending on the transceiver structure [151]. The standard proposes a calibration mechanism to compensate for the internal propagation time but some uncertainties still remain. Note that n can also contain communication and quantization errors as discussed next. Since the timestamps are distributed over the two nodes, they have to be brought together, which cannot be accomplished perfectly due to the limited communication resources. Furthermore, the abstract ranging counter, which assigns values to the timestamps, runs at a nominal 64 GHz in the standard, which causes some quantization effects.

We remark that the timestamps employed for ranging are different from the timestamps used in traditional clock synchronization protocols. Since the timestamps used here are recorded at the physical layer when the RMARKER leaves or arrives at the antenna, it excludes most of the conventional sources of uncertainty of message delivery delays in clock synchronization [27] [143] [158], including send time, access time, reception time and receive time, which are the main error sources in

Figure 7.1: An example of the TWR protocol

clock synchronization. The TOF for ranging, called the propagation time in clock synchronization, is one of the sources of uncertainty of message delivery delays, but it only contributes a little compared to other sources in traditional clock synchronization. Therefore, if these timestamps are used not only for ranging, but also for clock synchronization, a much better accuracy could be achieved than the existing clock synchronization protocols (see [27, 29, 143] and references therein). This kind of physical layer synchronization is also referred to as the pulse coupling method in [139].

Let us now focus on the security issues of the TWR protocol. The standard provides optional private ranging as a secure mode. The dynamic preamble selection and the encryption of the timestamp reports are used to facilitate the private ranging [75, 76]. However, the TWR protocol is vulnerable to an internal attack, which cannot be addressed by conventional cryptographic countermeasures. According to (7.1), the TOF t_0 depends not only on the timestamps T_{AR} and T_{AT} at node A, but also on the processing time τ_0 at node B. The dependence on the reliability and synchronization of two different nodes is a weak point of the TWR protocol. For example, assume node B is compromised and tries to cheat node A about its distance by tampering its processing time as τ_0'. Then, t_0 will be miscalculated, since node A is not aware of the attack. In the following, we adopt the same signal structure as in Fig. 7.3 and propose a new protocol which is immune to internal attacks.

7.3   Localization for quasi-synchronous networks

We again consider M anchor nodes and one target node, and would like to estimate the position of the target node. We recall that $\mathbf{X}_a$ indicates the known coordinates of the anchor nodes, and $\mathbf{x}$ denotes the unknown coordinates of the target node. In this section, we tackle quasi-synchronous networks, and we leave the fully
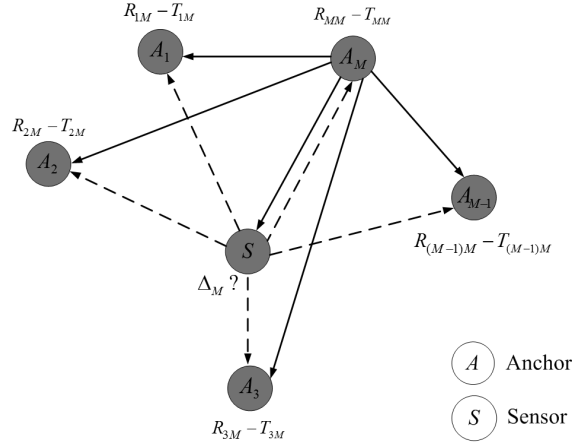
Figure 7.2: An illustration of the ATR protocol

asynchronous case to the next section.

7.3.1 System model

In quasi-synchronous networks, the target node clock runs freely, and the clock skews of all the anchors are equal to 1 or treated as 1. There are only clock offsets among all the anchors. The timing relation between the ith anchor clock $C_i(t)$ and the absolute time t can be described as [29]

$$C_i(t) = t + \theta_i, \ i = 1, 2, \ldots, M, \tag{7.2}$$

where θ_i is the unknown clock offset of $C_i(t)$ relative to the absolute clock. Moreover, the model for the target clock is given by

$$C_s(t) = \alpha_s t + \theta_s, \tag{7.3}$$

where α_s and θ_s denote the unknown clock skew and clock offset of the target node clock relative to the absolute clock.

A novel asymmetric trip ranging (ATR) protocol is shown in Fig. 7.2, which subsumes the protocol used in [135] as a special case. The ATR protocol makes all the other anchors listen to the ranging packets and record timestamps locally, when one anchor and the target node exchange their ranging packets. It can obtain more information than the TWR protocol, where all the other nodes are idle, when two nodes exchange their ranging packets. The ATR protocol starts with one of the anchors initiating the ranging request and recording a timestamp when its RMARKER departs, which can also be interpreted as the time when that anchor receives its own RMARKER without any delay. Without loss of generality,

we assume the Mth anchor initiates the ranging request, and we denote the time recorded at the Mth anchor as T_{MM}. Consequently, all the other anchors and the target node receive the ranging request and record their own timestamps T_{iM}, $i = 1, 2, \ldots, M-1$ and T_{SM}, respectively, as soon as they detect the RMARKER from the Mth anchor. The target node processes the ranging request and broadcasts a response. The departure time of the target RMARKER is recorded as R_{SM}, and we define $\tau = R_{SM} - T_{SM}$ as the true processing time of the target node. Each anchor in the network detects the broadcasted ranging response from the target node, and records its own timestamp for the arrival of the target RMARKER as R_{iM}, $i = 1, 2, \ldots, M$. If a compromised target node tampers its processing time as τ', or a target node reports τ' due to the clock skew or the internal delay, all the distance measurements would be decreased or enlarged by $c|\tau - \tau'|$ (where c is the speed of light), which would lead to a meaningless position estimate. Hence, this problem will be addressed later.

For the ith anchor node, the difference between R_{iM} and T_{iM} relates to the distance as

$$c(R_{iM} - T_{iM}) \quad = \quad d_i + d_M + \Delta_M - d_{iM} + n_{iM} - m_{iM}, \quad i = 1, 2, \ldots, M, \tag{7.4}$$

where $d_i = \|\mathbf{x}_i - \mathbf{x}\| = \sqrt{\|\mathbf{x}_i\|^2 - 2\mathbf{x}_i^T\mathbf{x} + \|\mathbf{x}\|^2}$ is the unknown distance between the ith anchor and the target node, $\Delta_M = c\tau$ is the unknown distance corresponding to the target node processing time in response to the Mth anchor, and $d_{iM} = \|\mathbf{x}_i - \mathbf{x}_M\|$ is the known distance between the ith and the Mth anchors. Furthermore, n_{iM} and m_{iM} denote the distance errors translated from the measurement errors of R_{iM} and T_{iM}, respectively, which are aggregate error terms, as we have discussed in Section 7.2. Note that the recordings of R_{iM} and T_{iM}, $i = 1, 2, \ldots, M-1$ are triggered by the received RMARKERs, and thus the same internal delays are involved, which are canceled out by making differences of timestamps recorded at the same node as indicated in (7.4). By making these differences, the clock offsets at the anchors are also canceled*. The situation is different for the Mth anchor, since it records T_{MM} and R_{MM} upon transmitting and receiving the RMARKERs, respectively. As a result, the internal delays of the transmission path and the receiving path are added up when computing $R_{MM} - T_{MM}$. Thus, we assume that the main part of the Mth anchor's internal delay is compensated beforehand as accomplished in [151]. But different from [151], compensation is not required for the other anchors in our scheme. Consequentially, n_{iM} and m_{iM} can be modeled as zero-mean random variables with variance σ_i^2 and σ_{iM}^2, respectively [157]. Note that the timestamps employed here are recorded at the physical layer, which are totally different from the conventional timestamps recorded at the MAC or other upper layers in clock synchronization,

---

*Note that this is different from the traditional TDOA approach, which requires synchronization among anchor nodes [9].

which have different error sources. A NLOS environment would introduce a biased LED error, and in that case n_{iM} or m_{iM} will have a non-zero mean. However, since this bias is not known, the only safe assumption is to view it as zero mean, or we assume that a calibration is carried out to remove it beforehand. More sophisticated error models such as the one in [149] can be considered in future work. Defining $\mathbf{u}_M = c[R_{1M}, R_{2M}, \ldots, R_{MM}]^T$, $\mathbf{v}_M = c[T_{1M}, T_{2M}, \ldots, T_{MM}]^T$, $\mathbf{d} = [d_1, d_2, \ldots, d_M]^T$, $\mathbf{g}_M = [d_{1M}, d_{2M}, \ldots, d_{(M-1)M}, 0]^T$, $\mathbf{n}_M = [n_{1M}, n_{2M}, \ldots, n_{MM}]^T$ and $\mathbf{m}_M = [m_{1M}, m_{2M}, \ldots, m_{MM}]^T$, we can write (7.4) in vector form as

$$\mathbf{u}_M - \mathbf{v}_M = \mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M + \mathbf{n}_M - \mathbf{m}_M. \qquad (7.5)$$

In order to be immune to an internal attack by the compromised target node or to incorrect timestamps due to the randomness of the target node clock, we do not employ the timestamp report from the target node, but only use it as a trigger at each anchor. We estimate the target position only based on the timestamps T_{iM} and $R_{iM}, i = 1, 2, \ldots, M$, recorded locally at the M anchors. Because we do not use the timestamps of the target node, its clock parameters, such as clock skew, clock offset and internal delay, do not have any impact on our scheme. This distinguishes our algorithm from others that use the timestamps of the target node, such as [28, 149, 151]. It is easy for the target node to cheat one anchor, but it is almost impossible to cheat all the anchors simultaneously. We remark that the cooperative positioning protocol proposed in [159] is similar to our ATR protocol. However, our method differs from [159] in several aspects: (i) we do not use the timestamps from the target node, and thus our method is more robust to unreliable timestamps; (ii) the target node processing time is unknown; and (iii) we propose low-complexity closed-form solutions for localization, instead of complex MLEs.

7.3.2 Localization algorithm

Since we do not use the timestamps from the target node, the clock parameters of the target node do not impact its position estimate. More specifically, we treat Δ_M (the distance corresponding to the target node processing time) in (7.5) as an unknown parameter. Note that (7.5) is a linear equation w.r.t. Δ_M, but it is a complicated nonlinear equation w.r.t. $\mathbf{x}$ due to d_M and $\mathbf{d}$. We are not interested in methods with a high computational complexity, such as the MLE which also requires the unknown noise pdf. Because of the low-cost and low-power constraints of a WSN, we explore low-complexity closed-form solutions for localization.

Since $\Delta_M \gg d_i$, Δ_M is a dominant term at the right hand side of (7.5). In order to extract useful distance information, we have to preprocess (7.5). Instead of choosing a reference anchor node as proposed in [135], we employ an orthogonal projection $\mathbf{P}$ onto the orthogonal complement of $\mathbf{1}_M$, which is given by $\mathbf{P} = \mathbf{I}_M -$

$\frac{1}{M}\mathbf{1}_M\mathbf{1}_M^T$. Since $\mathbf{P}\mathbf{1}_M = \mathbf{0}_M$, $\mathbf{P}$ can be used to eliminate the term $(d_M + \Delta_M)\mathbf{1}_M$ in (7.5). As a result, premultiplying both sides of (7.5) with $\mathbf{P}$, we obtain

$$\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M) = \mathbf{Pd} - \mathbf{Pg}_M + \mathbf{P}(\mathbf{n}_M - \mathbf{m}_M). \qquad (7.6)$$

Note that $\mathbf{Pd} = \mathbf{d} - \bar{d}\mathbf{1}_M$, where $\bar{d} = \frac{1}{M}\sum_{i=1}^M d_i$ is the unknown average of the distances between the target node and the anchors. Thus, (7.6) can be rewritten as

$$\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M) = \mathbf{d} - \bar{d}\mathbf{1}_M - \mathbf{Pg}_M + \mathbf{Pn}_M - \mathbf{Pm}_M. \qquad (7.7)$$

Keeping $\mathbf{d}$ on one side, moving the other terms to the other side, and making an element-wise multiplication, we achieve

$$\begin{aligned}
\boldsymbol{\psi}_a - 2\mathbf{X}_a^T\mathbf{x} + \|\mathbf{x}\|^2\mathbf{1}_M &= (\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M)) \odot (\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M)) \\
&\quad + \bar{d}^2\mathbf{1}_M + 2\bar{d}\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M) + \mathbf{n}_{rs}, \qquad (7.8)
\end{aligned}$$

where $\boldsymbol{\psi}_a = [\|\mathbf{x}_1\|^2, \|\mathbf{x}_2\|^2, \ldots, \|\mathbf{x}_M\|^2]^T$, and

$$\begin{aligned}
\mathbf{n}_{rs} \\
&= (\mathbf{P}(\mathbf{m}_M - \mathbf{n}_M)) \odot (\mathbf{P}(2(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M) + \mathbf{m}_M - \mathbf{n}_M)) + 2\bar{d}\mathbf{P}(\mathbf{m}_M - \mathbf{n}_M) \\
&= -(\mathbf{P}(\mathbf{n}_M - \mathbf{m}_M)) \odot (\mathbf{P}(\mathbf{n}_M - \mathbf{m}_M)) - 2\mathbf{d} \odot \mathbf{P}(\mathbf{n}_M - \mathbf{m}_M). \qquad (7.9)
\end{aligned}$$

Defining $\mathbf{Pm}_M = \mathbf{m}_M - \bar{m}_M\mathbf{1}_M$ and $\mathbf{Pn}_M = \mathbf{n}_M - \bar{n}_M\mathbf{1}_M$, where $\bar{m}_M = \frac{1}{M}\sum_{i=1}^M m_{iM}$ and $\bar{n}_M = \frac{1}{M}\sum_{i=1}^M n_{iM}$, we can write the entries of $\mathbf{n}_{rs}$ as

$$[\mathbf{n}_{rs}]_i = 2d_i(m_{iM} - \bar{m}_M - n_{iM} + \bar{n}_M) - (m_{iM} - \bar{m}_M - n_{iM} + \bar{n}_M)^2,$$
$$i = 1, 2, \ldots, M. \quad (7.10)$$

Recall that $E[n_{iM}] = 0$, $E[n_{iM}^2] = \sigma_i^2$ and $E[n_{iM}n_{jM}] = 0, i \neq j$, leading to $E[\bar{n}_M] = 0$, $E[\bar{n}_M^2] = \frac{1}{M^2}\sum_{i=1}^M \sigma_i^2$ and $E[\bar{n}_M n_{iM}] = \frac{1}{M}\sigma_i^2$. The stochastic properties of m_{iM} can be obtained in a similar way. Moreover, n_{iM} and $m_{iM}, i = 1, 2, \ldots, M$ are uncorrelated. As a result, the stochastic properties of $\mathbf{n}_{rs}$ are

given by

$$
E[[\mathbf{n}_{rs}]_i] = \frac{2-M}{M}(\sigma_{iM}^2 + \sigma_i^2) - \frac{1}{M^2}\sum_{k=1}^{M}(\sigma_{kM}^2 + \sigma_k^2) \approx 0, \tag{7.11}
$$

$$
\begin{aligned}
[\mathbf{\Sigma}_{rs}]_{i,j} &= E[[\mathbf{n}_{rs}]_i[\mathbf{n}_{rs}]_j] \\
&= E[(2d_i(m_{iM} - \bar{m}_M - n_{iM} + \bar{n}_M) - (m_{iM} - \bar{m}_M - n_{iM} + \bar{n}_M)^2) \\
&\quad \times (2d_j(m_{jM} - \bar{m}_M - n_{jM} + \bar{n}_M) - (m_{jM} - \bar{m}_M - n_{jM} + \bar{n}_M)^2)], \\
&\approx 4d_i d_j E[(m_{iM} - \bar{m}_M - n_{iM} + \bar{n}_M)(m_{jM} - \bar{m}_M - n_{jM} + \bar{n}_M)], \\
&\approx \begin{cases}
4d_i^2\left(\frac{M-2}{M}(\sigma_{iM}^2 + \sigma_i^2) + \frac{1}{M^2}\sum_{k=1}^{M}(\sigma_{kM}^2 + \sigma_k^2)\right) \\
\quad i = j, \\
4d_i d_j\left(\frac{1}{M^2}\sum_{k=1}^{M}(\sigma_k^2 + \sigma_{kM}^2) - \frac{1}{M}(\sigma_i^2 + \sigma_j^2 + \sigma_{iM}^2 + \sigma_{jM}^2)\right) \\
\quad i \neq j,
\end{cases}
\end{aligned}
$$
$$\tag{7.12}$$

where we ignore the higher order noise terms to obtain (7.12) and assume $E[[\mathbf{n}_{rs}]_i] \approx 0$ under the condition of sufficiently small measurement errors. Note that the noise covariance matrix $\mathbf{\Sigma}_{rs}$ depends on the unknown $\mathbf{d}$.

As (7.8) is still a nonlinear equation w.r.t. $\mathbf{x}$, we make again use of the orthogonal projection $\mathbf{P}$ to eliminate the terms $\|\mathbf{x}\|^2$ and $\bar{d}^2$ in (7.8). By premultiplying both sides of (7.8) with $\mathbf{P}$ and rearranging the terms, we arrive at

$$
\begin{aligned}
\mathbf{P}\boldsymbol{\psi}_a &- \mathbf{P}((\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M)) \odot (\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M))) \\
&= 2\mathbf{P}\mathbf{X}_a^T\mathbf{x} + 2\bar{d}\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M) + \mathbf{P}\mathbf{n}_{rs}.
\end{aligned}
$$
$$\tag{7.13}$$

As a result, (7.13) becomes a linear equation w.r.t. both $\mathbf{x}$ and $\bar{d}$. Defining $\boldsymbol{\phi}_q = \boldsymbol{\psi}_a - ((\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M)) \odot (\mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M)))$, $\mathbf{H}_q = 2\left[\mathbf{X}_a^T, \ \mathbf{P}(\mathbf{u}_M - \mathbf{v}_M + \mathbf{g}_M)\right]$, and $\mathbf{s}_q = [\mathbf{x}^T, \ \bar{d}]^T$, we can finally rewrite (7.13) as

$$
\mathbf{P}\boldsymbol{\phi}_q = \mathbf{P}\mathbf{H}_q\mathbf{s}_q + \mathbf{P}\mathbf{n}_{rs}. \tag{7.14}
$$

We can find the LS and WLS solutions for (7.14) as

$$
\hat{\mathbf{s}}_q = (\mathbf{H}_q^T\mathbf{P}\mathbf{H}_q)^{-1}\mathbf{H}_q^T\mathbf{P}\boldsymbol{\phi}_q, \tag{7.15}
$$

and

$$
\hat{\mathbf{s}}_q = (\mathbf{H}_q^T\mathbf{P}\mathbf{W}\mathbf{P}\mathbf{H}_q)^{-1}\mathbf{H}_q^T\mathbf{P}\mathbf{W}\mathbf{P}\boldsymbol{\phi}_q, \tag{7.16}
$$

respectively, where $\mathbf{W}$ is a weighting matrix. The optimal weighting matrix $\mathbf{W}_o$

is given by

$$\mathbf{W}_o = (\mathbf{P}\boldsymbol{\Sigma}_{rs}\mathbf{P})^\dagger, \tag{7.17}$$

where we use the pseudo inverse because the $M \times M$ projection matrix $\mathbf{P}$ has rank $M - 1$. Furthermore, $\mathbf{PH}_q$ should be a full rank tall matrix. Thus, the number of anchors M should be no less than $l + 3$ to facilitate two projection operations, which indicates that we need at least five anchors to estimate the target position on a plane. Since $\mathbf{W}_o$ depends on the unknown $\mathbf{d}$, we can update it iteratively. Consequently, the iterative WLS is summarized as follows

1. Initialize $\mathbf{W}$ using the estimate of $\mathbf{d}$ based on the LS estimate of $\mathbf{x}$;

2. Estimate $\hat{\mathbf{s}}_q$ using (7.16);

3. Construct $\mathbf{W}$ using (7.17), where $\boldsymbol{\Sigma}_{rs}$ is computed using $\hat{\mathbf{s}}_s$;

4. Repeat Steps 2) and 3) until no obvious improvement of the cost function $(\boldsymbol{\phi}_q - \mathbf{H}_q\mathbf{s}_q)^T \mathbf{PWP}(\boldsymbol{\phi}_q - \mathbf{H}_q\mathbf{s}_q)$ is observed.

An estimate of $\mathbf{x}$ is finally given by

$$\hat{\mathbf{x}} = [\mathbf{I}_l \quad \mathbf{0}_l]\hat{\mathbf{s}}_q. \tag{7.18}$$

We remark that the estimator (7.15) (or (7.16)) is equivalent to the unconstrained LS (or WLS) estimator to obtain $\mathbf{x}, \bar{d}$ and $\bar{d}^2 - \|\mathbf{x}\|^2$ all together as discussed in [123]. We may even improve the estimation performance by exploring the relations among $\mathbf{x}$, $\bar{d}$ and $\bar{d}^2 - \|\mathbf{x}\|^2$ as constraints. Constrained LS (CLS) and weighted CLS estimators can be derived as in [6, 11]. However, it is extremely difficult to take the relation between $\mathbf{x}$ and $\bar{d}$ into account, since it is highly non-linear.

The distance Δ_M corresponding to the target node processing time in response to the Mth anchor can be estimated as

$$\hat{\Delta}_M = \mathbf{1}_M^T(\mathbf{u}_M - \mathbf{v}_M - \hat{\mathbf{d}} + \mathbf{g}_M) - \hat{d}_M, \tag{7.19}$$

where $\hat{d}_i = \|\hat{\mathbf{x}} - \mathbf{x}_i\|$, $i = 1, 2, \ldots, M$ are the distance estimates between the target node and the anchors based on $\hat{\mathbf{x}}$. We remark that there are mathematical similarities between our data model (7.5) and the data model in [151], if we regard $d_M + \Delta_M$ in (7.5) as an unknown internal delay. However, we employ a novel ATR protocol and estimate the parameters in a different way.

7.4 Localization for fully asynchronous networks

In this section, we relax all the synchronization constraints on the anchors and the target node. There are not only clock offsets, but also clock skews among all

the anchors and the target node. We use the same clock model (7.3) here for the target node to indicate its clock skew α_s and clock offset θ_s. The anchor clock model $C_i(t)$ is now given by

$$C_i(t) \quad = \quad \alpha_i t + \theta_i, \ \ i = 1, 2, \ldots, M, \tag{7.20}$$

where α_i denotes the unknown clock skew of $C_i(t)$ relative to the absolute clock, and θ_i again is the unknown clock offset. Thus, the relations between the clocks $C_i(t)$ and $C_j(t)$ are given by,

$$C_i(t) \quad = \quad \frac{\alpha_i}{\alpha_j} C_j(t) - \frac{\alpha_i}{\alpha_j} \theta_j + \theta_i. \tag{7.21}$$

Applying the same ATR protocol as in Section 7.3.1, we obtain T_{iM} and $R_{iM}, i = 1, 2, \ldots, M$, which are in total $2M$ timestamps. Fig. 7.3 shows an example of a transmitted ranging packet and a received ranging packet. The time intervals in Fig. 7.3 are measured by the local clocks of the nodes. The length of the preamble is defined as T_p. Since the ranging packet is generated by the Mth anchor, it generates the preamble of length T_p relative to its own clock. The ith anchor regards the length of the received preamble as $\alpha_i T_p / \alpha_M$ due to the relative clock skew. Therefore, the difference between T_{iM} and T_{MM} is not only related to the TOF between the anchors (measured by the ith anchor as $\alpha_i d_{iM}/c$), but also to the relative clock drift over the whole preamble. However, the relations between R_{iM} and T_{iM} the timestamps recorded at the same anchor are relatively simple. By making differences of the timestamps from the same anchor, the clock offsets are again canceled out. As a result, the difference between R_{iM} and T_{iM} can be described as

$$\frac{c}{\alpha_i}(R_{iM} - T_{iM}) \quad = \quad d_i + d_M + \Delta_M - d_{iM} + \frac{n_{iM}}{\alpha_i} - \frac{m_{iM}}{\alpha_i}, \quad i = 1, 2, \ldots, M. \tag{7.22}$$

There is no impact of the relative clock drift over the whole preamble in (7.22). Note that the error terms n_{iM} and m_{iM} are also influenced by the clock skew of the ith anchor. We remark that the target clock does not have any impact on (7.22), but the influence of the asynchronous anchors remains. This again confirms that ignoring the timestamps from the target node can thwart internal attacks.

7.4.1   Localization approach I

In (7.22), there are M equations and $M + l + 1$ unknown parameters in total. It is clear that there are not enough observation data (timestamps) to estimate all the parameters, if the ranging procedure is only executed once. To solve that problem, we propose two different approaches. In a power-hungry WSN, data
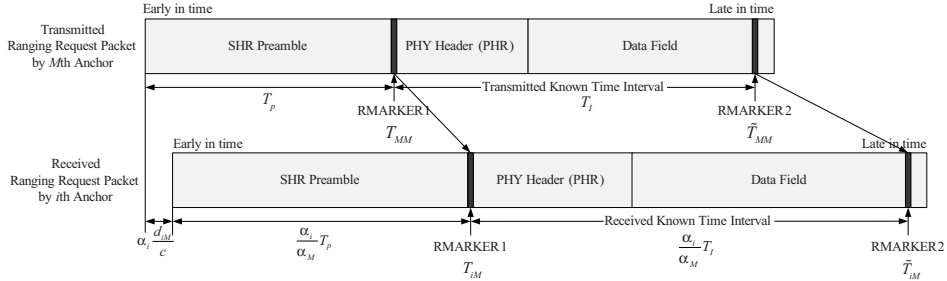
Figure 7.3: An example of a transmitted ranging packet and a received ranging packet

communications could consume much more power than data processing [160], and the communication load should be reduced as much as possible. Therefore, we propose localization approach I in order to achieve the minimum communication load. On the other hand, if we can afford a bit more communication load, the ranging procedure can be repeated several times in order to obtain sufficient observations for accurate estimation, which is referred as localization approach II. We will first introduce localization approach I in this subsection, and explain localization approach II in the next subsection.

System model

As we mentioned before, in a power-hungry WSN, the communication load should be reduced as much as possible. Therefore, we propose localization approach I to add another RMARKER in the ranging packet to facilitate the clock synchronization in order to achieve a small communication load. The new ranging packet not only includes RMARKER1 in the PHY header (PHR), but also has another RMARKER2 as the first pulse in the last symbol of the data field as shown in Fig. 7.3. Although in the standard, there is only one RMARKER, we can literally define the second one without changing anything in the ranging packet. We further make use of the prior knowledge about the length of the ranging packet, and thus the time interval T_I between RMARKER1 and RMARKER2 is known to all the anchors. The length of the received ranging packet can be extended or reduced due to the relative clock skew of the receiver clock, and thus yields the parameter bearing the information of the relative clock skew. This parameter is different from the ones used in [151, 157], where ranging packets have to be consecutively or periodically transmitted. By making use of the known length of the ranging packet [157] or the known transmission period [151], they also obtain parameters bearing the relative clock skew information. However, our scheme has a smaller communication load than theirs. In Fig. 7.3, we show an example of an extended ranging packet, where the clock skew difference between the receiver and the transmitter is positive. We remark that T_I should be long enough to

observe sufficient clock drift, which means that after T_I the clock drift should be larger than the resolution of the TOA estimation. According to the standard, the time interval T_I could be several milliseconds, e.g. 5 ms. With a typical value of the relative clock skew of 40 ppm, we will observe 200 ns clock difference after 5 ms, whereas the resolution of TOA estimation using an UWB signal with a 1 GHz bandwidth can reach several nanoseconds [1]. Therefore, using a standard ranging packet, it is sufficient to estimate the drift between RMARKER1 and RMARKER2. Consequently, each anchor records two timestamps, when it receives a ranging packet during the procedure. The first set of timestamps is still represented by T_{iM} and R_{iM}, while the second one is denoted by $\widetilde{T}_{iM}$ and $\widetilde{R}_{iM}$. Their relations are summarized as

$$\frac{c}{\alpha_i}(\widetilde{T}_{iM} - T_{iM}) = \frac{cT_I}{\alpha_M} + \frac{\widetilde{m}_{iM}}{\alpha_i} - \frac{m_{iM}}{\alpha_i}, \ i = 1, 2, \ldots, M, \qquad (7.23)$$

$$\frac{c}{\alpha_i}(\widetilde{R}_{iM} - R_{iM}) = \frac{cT_I}{\alpha_s} + \frac{\widetilde{n}_{iM}}{\alpha_i} - \frac{n_{iM}}{\alpha_i}, \ i = 1, 2, \ldots, M, \qquad (7.24)$$

where $\widetilde{m}_{iM}$ and $\widetilde{n}_{iM}$ are the corresponding distance errors due to the measurement errors of $\widetilde{T}_{iM}$ and $\widetilde{R}_{iM}$, respectively. We assume that the variances of the measurement errors for RMARKER1 and RMARKER2 in the same packet are the same, and thus $\widetilde{m}_{iM}$ and $\widetilde{n}_{iM}$ are also modeled as zero-mean random variables with variances σ_{iM}^2 and σ_i^2, respectively. Note that the clock skew of the target α_s in (7.24) influences the time difference T_I between RMARKER1 and RMARKER2 of the ranging response, which is generated by the target node. We remark that the other clock parameters of the target node, i.e. the clock offset and the internal delay, do not influence our scheme, since we still do not use the timestamps from the target node.

We can now write (7.22), (7.23) and (7.24) in vector form as

$$\mathbf{F}_1\boldsymbol{\gamma} = \frac{cT_I}{\alpha_M}\mathbf{1}_M + \mathrm{diag}(\boldsymbol{\gamma})(\widetilde{\mathbf{m}}_M - \mathbf{m}_M), \qquad (7.25)$$

$$\mathbf{F}_2\boldsymbol{\gamma} = \frac{cT_I}{\alpha_s}\mathbf{1}_M + \mathrm{diag}(\boldsymbol{\gamma})(\widetilde{\mathbf{n}}_M - \mathbf{n}_M), \qquad (7.26)$$

$$\mathbf{F}_3\boldsymbol{\gamma} = \mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M + \mathrm{diag}(\boldsymbol{\gamma})(\mathbf{n}_M - \mathbf{m}_M), \qquad (7.27)$$

where $\boldsymbol{\gamma} = [1/\alpha_1 \quad 1/\alpha_2 \quad \ldots \quad 1/\alpha_M]^T$, $\mathbf{F}_1 = \mathrm{diag}(\widetilde{\mathbf{v}}_M - \mathbf{v}_M)$, $\mathbf{F}_2 = \mathrm{diag}(\widetilde{\mathbf{u}}_M - \mathbf{u}_M)$, $\mathbf{F}_3 = \mathrm{diag}(\mathbf{u}_M - \mathbf{v}_M)$, $\widetilde{\mathbf{v}}_M = c[\widetilde{T}_{1M}, \widetilde{T}_{2M}, \ldots, \widetilde{T}_{MM}]^T$, $\widetilde{\mathbf{u}}_M = c[\widetilde{R}_{1M}, \widetilde{R}_{2M}, \ldots, \widetilde{R}_{MM}]^T$, $\widetilde{\mathbf{m}}_M = [\widetilde{m}_{1M}, \widetilde{m}_{2M}, \ldots, \widetilde{m}_{MM}]^T$ and $\widetilde{\mathbf{n}}_M = [\widetilde{n}_{1M}, \widetilde{n}_{2M}, \ldots, \widetilde{n}_{MM}]^T$. Recall that $\mathbf{u}_M = c[R_{1M}, R_{2M}, \ldots, R_{MM}]^T$ and $\mathbf{v}_M = c[T_{1M}, T_{2M}, \ldots, T_{MM}]^T$. Our goal is to estimate $\mathbf{x}$, $\boldsymbol{\gamma}$ and Δ_M based on (7.25) - (7.27). Note that the last equation in (7.25) does not offer any useful information, since $1/\alpha_M$ is at both sides of the equation. Thus, we collect the

$3M - 1$ equations related to $\boldsymbol{\gamma}$ and write them in vector form as

$$\mathbf{F}\boldsymbol{\gamma} \;=\; \begin{bmatrix} \mathbf{0}_{M-1} \\ \dfrac{1}{\alpha_s}cT_I\mathbf{1}_M \\ \mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M \end{bmatrix} + \begin{bmatrix} \mathbf{Z}\mathrm{diag}(\boldsymbol{\gamma})(\widetilde{\mathbf{m}}_M - \mathbf{m}_M) \\ \mathrm{diag}(\boldsymbol{\gamma})(\widetilde{\mathbf{n}}_M - \mathbf{n}_M) \\ \mathrm{diag}(\boldsymbol{\gamma})(\mathbf{n}_M - \mathbf{m}_M) \end{bmatrix},$$

$$(7.28)$$

where $\mathbf{F} = [\mathbf{F}_1^T\mathbf{Z}^T - \mathbf{E}_1^T,\ \mathbf{F}_2^T,\ \mathbf{F}_3^T]^T$, $\mathbf{E}_1 = [\mathbf{0}_{(M-1)\times(M-1)},\ cT_I\mathbf{1}_{M-1}]$ and $\mathbf{Z} = [\mathbf{I}_{M-1},\ \mathbf{0}_{M-1}]$.

Localization algorithm

From now on, we ignore the error terms in the analysis for simplicity. The localization algorithms based on (7.25)-(7.27) are also naturally immune to the unreliable timestamps from the target node, and robust to the randomness of the target node clock, since we do not use the timestamps from the target node. We would like to investigate low-complexity localization methods. Although the data model (7.28) is a complicated nonlinear equation w.r.t. $\mathbf{x}$, it is linear w.r.t. to $\boldsymbol{\gamma}$. We can first estimate $\boldsymbol{\gamma}$ as a function of $\mathbf{x}$ and $1/\alpha_s$ based on (7.28), and then estimate $\mathbf{x}$ based on (7.27) by plugging in the estimate of $\boldsymbol{\gamma}$. From (7.28), the LS estimate of $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}} = (\mathbf{F}^T\mathbf{F})^{-1}\Big(\frac{1}{\alpha_s}cT_I(\widetilde{\mathbf{u}}_M - \mathbf{u}_M) + \mathbf{F}_3^T(\mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M)\Big). \quad (7.29)$$

Plugging (7.29) into (7.27), and rearranging the terms, we achieve

$$\mathbf{K}_1(\mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M) = \frac{1}{\alpha_s}cT_I\mathbf{F}_3(\mathbf{F}^T\mathbf{F})^{-1}(\widetilde{\mathbf{u}}_M - \mathbf{u}_M), \quad (7.30)$$

where $\mathbf{K}_1 = \mathbf{I}_M - \mathbf{F}_3(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}_3^T$, which is full rank, and whose inverse is explored in Appendix 7.A. Premultiplying both sides of (7.30) with $\mathbf{K}_1^{-1}$, we arrive at

$$\mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M \;=\; \frac{1}{\alpha_s}\mathbf{h}, \quad (7.31)$$

where $\mathbf{h} = cT_I\mathbf{K}_1^{-1}\mathbf{F}_3(\mathbf{F}^T\mathbf{F})^{-1}(\widetilde{\mathbf{u}}_M - \mathbf{u}_M)$. We remark that at this point an MLE can be derived to jointly estimate $\mathbf{x}$, Δ_M and $1/\alpha_s$ based on (7.31) via exhaustive search. It needs at least four anchors to locate a target node on a plane in this case. However, it has a high computational complexity. Thus, we continue to investigate low-complexity closed-form solutions. Applying $\mathbf{P}$ to get rid of $(d_M + \Delta_M)\mathbf{1}_M$, recalling that $\mathbf{P}\mathbf{d} = \mathbf{d} - \bar{d}\mathbf{1}_M$, and moving $\mathbf{d}$ to one side

and the other terms to the other side, we obtain

$$\mathbf{d} \ = \ \frac{1}{\alpha_s}\mathbf{Ph} + \bar{d}\mathbf{1}_M + \mathbf{Pg}_M. \tag{7.32}$$

Executing element-wise multiplication and rearranging the equation, we obtain the linear equation

$$\boldsymbol{\phi} \ = \ \mathbf{Hs}, \tag{7.33}$$

where $\boldsymbol{\phi} = \boldsymbol{\psi}_a - (\mathbf{Pg}_M) \odot (\mathbf{Pg}_M)$, $\mathbf{s} = [\mathbf{x}^T,\ \bar{d},\ \frac{1}{\alpha_s},\ (\bar{d}^2 - \|\mathbf{x}\|^2),\ \frac{1}{\alpha_s^2},\ 2\frac{\bar{d}}{\alpha_s}]^T$, and $\mathbf{H} = [2\mathbf{X}_a^T,\ 2\mathbf{Pg}_M,\ 2(\mathbf{Ph}) \odot (\mathbf{Pg}_M),\ \mathbf{1}_M,\ (\mathbf{Ph}) \odot (\mathbf{Ph}),\ \mathbf{Ph}]$. However, as investigated in Appendix 7.B, $\mathbf{H}$ is always rank-deficient in the noiseless case or with sufficiently small noise. Hence, it is impossible to have a unique estimate of $\mathbf{s}$ based on (7.33). We can also interpret this problem from another point of view. Our method is equivalent to first jointly estimating $\boldsymbol{\gamma}$, $\mathbf{d}$, $1/\alpha_s$, and $d_M + \Delta_M$ (in total $2M + 2$ parameters) based on (7.25)-(7.27), then plugging the estimate of $\boldsymbol{\gamma}$ into (7.27), linearizing the equations w.r.t. $\mathbf{x}$, and finally estimating $\mathbf{x}$. But (7.25) and (7.26) are linearly dependent in the noiseless case or with sufficiently small noise, and there are only $2M$ independent equations in (7.25)-(7.27). Therefore, we cannot estimate $\boldsymbol{\gamma}$ without ambiguities, and that is why (7.33) does not have a unique solution.

However, let us take some practical issues into account to solve this problem. The clock skew of the ith anchor relative to the absolute time is $\alpha_i - 1$, which is in the order of several tens of ppm (10^{-6}). The typical range of $\alpha_i - 1$ is from 2 ppm to 80 ppm according to the standard, which means that α_i is in the range of 0.99992 to 1.00008, and also α_s is in the same range. Thus, we can make a first-order Taylor expansion of $1/\alpha_s^2$ as a function of $1/\alpha_s$ around 1 by ignoring the higher order terms as $1/\alpha_s^2 \approx 2/\alpha_s - 1$. We can plug it into (7.33), rearrange the terms, and then obtain an equation w.r.t. $\mathbf{x}$, $\bar{d}$, $\bar{d}/\alpha_s$, $1/\alpha_s$ and $\bar{d}^2 - \|\mathbf{x}\|^2$ (in total $l + 4$ unknowns). However, if we make further use of the prior knowledge that α_s is very close to 1, we can obtain an equation with even fewer unknowns, leading to a better estimation performance. Thus, we can further assume that $\bar{d}/\alpha_s \approx \bar{d}$, and plug this together with $1/\alpha_s^2 \approx 2/\alpha_s - 1$ into (7.33), which leads to

$$\boldsymbol{\phi}_d \ = \ \mathbf{H}_d\mathbf{s}_d, \tag{7.34}$$

where $\boldsymbol{\phi}_d = \boldsymbol{\phi} + (\mathbf{Ph}) \odot (\mathbf{Ph})$, $\mathbf{s}_d = [\mathbf{x}^T,\ \bar{d},\ \frac{1}{\alpha_s},\ (\bar{d}^2 - \|\mathbf{x}\|^2)]^T$ and $\mathbf{H}_d = [2\mathbf{X}_a^T,\ 2\mathbf{P}(\mathbf{g}_M + \mathbf{h}),\ 2\mathbf{Ph} \odot (\mathbf{Pg}_M + \mathbf{Ph}),\ \mathbf{1}_M]$. Note that there are only $l + 3$ unknowns in (7.34). The LS estimate of $\mathbf{s}_d$ is then given by

$$\hat{\mathbf{s}}_d \ = \ (\mathbf{H}_d^T\mathbf{H}_d)^{-1}\mathbf{H}_d^T\boldsymbol{\phi}. \tag{7.35}$$

Note that the rank of $\mathbf{H}_d$ should be $l + 3$ in order to estimate all the parameters, which indicates $M - 1 \geq l + 3$. It needs at least six anchors to estimate all the parameters in a plane.

Now that we can simplify the problem by considering practical issues, let us revisit the data model (7.32). Since α_s is very close to 1, we may assume $\frac{1}{\alpha_s}\mathbf{Ph} \approx \mathbf{Ph}$ by ignoring the effect of $1/\alpha_s$, which means we treat α_s as 1, although it may not be exactly equal to 1. Note that this approximation can be improved by first applying conventional synchronization methods to the target node. We can then rewrite (7.32) as

$$\mathbf{d} \approx \mathbf{Ph} + \bar{d}\mathbf{1}_M + \mathbf{Pg}_M. \tag{7.36}$$

Sequentially, after element-wise multiplication and moving terms, we arrive at

$$\boldsymbol{\phi}_a = \mathbf{H}_a\mathbf{s}_a, \tag{7.37}$$

where $\boldsymbol{\phi}_a = \boldsymbol{\psi}_a - (\mathbf{P}(\mathbf{g}_M + \mathbf{h})) \odot (\mathbf{P}(\mathbf{g}_M + \mathbf{h}))$, $\mathbf{s}_a = [\mathbf{x}^T, \bar{d}, \bar{d}^2 - \|\mathbf{x}\|^2]^T$, and $\mathbf{H}_a = [2\mathbf{X}_a^T, 2\mathbf{P}(\mathbf{g}_M + \mathbf{h}), \mathbf{1}_M]$. The LS estimate of $\mathbf{s}_a$ is then given by

$$\hat{\mathbf{s}}_a = (\mathbf{H}_a^T\mathbf{H}_a)^{-1}\mathbf{H}_a^T\boldsymbol{\phi}_a. \tag{7.38}$$

The rank of $\mathbf{H}_a$ should be $l + 2$ in order to estimate all the parameters, which indicates $M - 1 \geq l + 2$. It needs at least five anchors to estimate all the parameters in a plane. As a result, taking such practical issues into account can dramatically simplify the problem. Moreover, we remark that the data model (7.36) is similar to the data model (7.7) in Section 7.3.2. Thus, the localization algorithms in Section 7.3.2 for quasi-synchronous networks can also be applied here to estimate $\mathbf{x}$.

Since it is always complicated to estimate $\boldsymbol{\gamma}$ and $\mathbf{x}$ jointly, we can also resort to simple solutions to decouple the synchronization and the localization. Defining $\boldsymbol{\beta}_M = \alpha_M[1/\alpha_1, 1/\alpha_2, \ldots, 1/\alpha_{M-1}]^T$ of length $M-1$, and $\widetilde{\boldsymbol{\beta}}_M = [\boldsymbol{\beta}^T, \alpha_M/\alpha_s]^T$ of length M, we can first estimate $\widetilde{\boldsymbol{\beta}}_M$ based on (7.25) and (7.26), which means that we first make use of T_I between the two RMARKERs to **C**alibrate the **C**lock **S**kews, and then **E**stimate the **N**ode **P**osition (CCS-ENP). As α_M is tightly coupled with the other clock skews, we can only estimate their ratios. We combine (7.25) and (7.26) ignoring the noise terms as

$$\mathbf{F}_\beta\widetilde{\boldsymbol{\beta}}_M = \mathbf{b}_\beta, \tag{7.39}$$

where $\mathbf{b}_\beta = [cT_I\mathbf{1}_{M-1}^T, \mathbf{0}_{M-1}^T, \widetilde{u}_M - u_M]^T$, and

$$\mathbf{F}_\beta = \begin{bmatrix} \mathbf{Z}\mathbf{F}_1\mathbf{Z}^T & \mathbf{0}_{M-1} \\ \mathbf{Z}\mathbf{F}_2\mathbf{Z}^T & -cT_I\mathbf{1}_{M-1} \\ \mathbf{0}_{M-1}^T & -cT_I \end{bmatrix}. \tag{7.40}$$

Consequently, the LS estimates of $\widetilde{\boldsymbol{\beta}}_M$ and $\boldsymbol{\beta}_M$ are given by

$$\hat{\widetilde{\boldsymbol{\beta}}}_M = (\mathbf{F}_\beta^T \mathbf{F}_\beta)^{-1} \mathbf{F}_\beta^T \mathbf{b}_\beta, \qquad (7.41)$$

$$\hat{\boldsymbol{\beta}}_M = \mathbf{Z}\hat{\widetilde{\boldsymbol{\beta}}}_M, \qquad (7.42)$$

respectively. Since only (7.27) is related to $\mathbf{x}$, we rewrite it as an equation in $\boldsymbol{\beta}_M$ and $\mathbf{x}$ without noise terms

$$\mathbf{F}_3 \boldsymbol{\beta}_a = \alpha_M(\mathbf{d} - \mathbf{g}_M + (d_M + \Delta_M)\mathbf{1}_M), \qquad (7.43)$$

where $\boldsymbol{\beta}_a = [\boldsymbol{\beta}_M^T, \ 1]^T$. Due to the fact that $\alpha_M \approx 1$, we can write $\alpha_M \mathbf{d} \approx \mathbf{d}$ and $\alpha_M \mathbf{g}_M \approx \mathbf{g}_M$. Thus, we can rewrite (7.43) as

$$\mathbf{F}_3 \boldsymbol{\beta}_a \approx \mathbf{d} + (d_M + \alpha_M \Delta_M)\mathbf{1}_M - \mathbf{g}_M, \qquad (7.44)$$

where α_M is tightly coupled with Δ_M. Again, we can improve this approximation by first synchronizing the Mth anchor in order to improve the localization performance. Plugging $\hat{\boldsymbol{\beta}}_M$ into (7.44), we observe that it is equivalent to the position estimation based on (7.5). Therefore, the time-based localization algorithms in Section 7.3.2 for quasi-synchronous networks can again be applied here to estimate $\mathbf{x}$. For brevity, we do not repeat the algorithms here. We remark that employing this separate method, we also need five anchors to calibrate the clock skews and estimate the target position in a plane. As an extension, even if some of the anchors are manipulated by attackers to report misinformation, we could still combine our approach with the methods in [118] or [156] to mitigate the influence of outliers. This is left for future work.

7.4.2 Localization approach II

System Model

If we may relax the communication load constraint, we can repeat the ATR protocol using different anchors as initiators to obtain additional information. Note that if an additional anchor plays the role of initiator, we obtain M new equations and one extra unknown distance corresponding to the processing time. Thus, we generalize (7.27), which is based on the Mth anchor to the following equation, where the jth anchor is used as initiator:

$$\text{diag}(\mathbf{u}_j - \mathbf{v}_j)\boldsymbol{\gamma} = \mathbf{d} + (d_j + \Delta_j)\mathbf{1}_M - \mathbf{g}_j + \text{diag}(\boldsymbol{\gamma})(\mathbf{n}_j - \mathbf{m}_j), \qquad (7.45)$$

where Δ_j is the unknown distance corresponding to the processing time of the target node formulating a response to the jth anchor. Moreover, $\mathbf{u}_j = c[R_{1j}, R_{2j}, \ldots, R_{Mj}]^T$, $\mathbf{v}_j = c[T_{1j}, T_{2j}, \ldots, T_{Mj}]^T$, $\mathbf{g}_j = [d_{1j}, d_{2j}, \ldots, d_{Mj}]^T$, $\mathbf{n}_j = [n_{1j}, n_{2j}, \ldots, n_{Mj}]^T$, and $\mathbf{m}_j = [m_{1j}, m_{2j}, \ldots, m_{Mj}]^T$.

Assuming that n anchors play the role of initiator, we have to fulfill the condition $nM \geq M + l + n$ in order to obtain enough equations to estimate all the parameters, where $M \geq n > 0$. It is possible that only a subset of anchors plays the role of initiator. However, since we are interested in the minimum number of anchors required for this approach, we take $n = M$, which means that all anchors participate. The minimum value of M is then given by $M_{\min} = \min\{M \in \{1, 2, \ldots\} | M^2 - 2M \geq l\}$, for instance, when $l = 2$, $M_{\min} = 3$. From now on, we ignore the error terms for simplicity, and assume $n = M$. We then obtain in total M^2 equations and can write them in vector form as follows

$$\mathbf{A}\boldsymbol{\gamma} = (\mathbf{B} + \mathbf{C})\mathbf{d} + \mathbf{C}\boldsymbol{\Delta} - \mathbf{g}, \tag{7.46}$$

where $\mathbf{A} = [\text{diag}(\mathbf{u}_1 - \mathbf{v}_1),\ \text{diag}(\mathbf{u}_2 - \mathbf{v}_2),\ \ldots,\ \text{diag}(\mathbf{u}_M - \mathbf{v}_M)]^T$, $\mathbf{B} = \mathbf{1}_M \otimes \mathbf{I}_M$, $\mathbf{C} = \mathbf{I}_M \otimes \mathbf{1}_M$, $\boldsymbol{\Delta} = [\Delta_1,\ \Delta_2,\ \ldots,\ \Delta_M]^T$, $\mathbf{g} = [\mathbf{g}_1^T,\ \mathbf{g}_2^T,\ \ldots,\ \mathbf{g}_M^T]^T$, and $\otimes$ denotes Kronecker product. We remark that the target node is required to use different processing times in response to different anchors. If $\Delta_i = \Delta_j$, we obtain the relation $(R_{ij} - T_{ij})/(R_{ji} - T_{ji}) = \alpha_i/\alpha_j$ without error terms, and it is only possible to estimate the relative clock skew. In that case, the equations (7.46) are not independent. Therefore, we assume that $\Delta_i \neq \Delta_j, i, j \in \{1, 2, \ldots, M\}$.

Localization Algorithm

We would like to estimate $\mathbf{x}$, $\boldsymbol{\gamma}$ and $\boldsymbol{\Delta}$, in total $2M + l$ unknown parameters based on (7.46), which is a complicated nonlinear equation w.r.t. $\mathbf{x}$. When we ignore the relations among the distances in $\mathbf{d}$ and regard them as independent unknowns, (7.46) is linear w.r.t. $\boldsymbol{\gamma}$, $\mathbf{d}$ and $\boldsymbol{\Delta}$. Therefore, we propose a two-step approach. We first jointly estimate $\boldsymbol{\gamma}$, $\mathbf{d}$ and $\boldsymbol{\Delta}$, obtaining a unique estimate for $\boldsymbol{\gamma}$ but ambiguous estimates of $\mathbf{d}$ and $\boldsymbol{\Delta}$. Secondly, we plug the estimate of $\boldsymbol{\gamma}$ into (7.46), linearize the equation w.r.t $\mathbf{x}$ via mathematical manipulations and then estimate $\mathbf{x}$. We remark that although our approach is accomplished in two steps, the first step is still a joint approach to estimate $\boldsymbol{\gamma}$, $\mathbf{d}$ and $\boldsymbol{\Delta}$, and in the second step the estimate of $\boldsymbol{\gamma}$ is used and the relation between $\mathbf{d}$ and $\mathbf{x}$ is explored. Furthermore, we use the same set of measurements to obtain all the estimates, and thus our method yields a joint synchronization and localization approach.

In the first step, there are $3M$ unknowns ($\boldsymbol{\gamma}$, $\mathbf{d}$ and $\boldsymbol{\Delta}$) and M^2 equations. Thus, $M_{\min} = 3$ is still valid. However, we note that $[\mathbf{B} + \mathbf{C},\ \mathbf{C}]$ is rank deficient because of the common basis of $\mathbf{B}$ and $\mathbf{C}$. Therefore, we can only jointly estimate $\mathbf{d}$ and $\boldsymbol{\Delta}$ with ambiguities. Since we are only interested in the result for $\boldsymbol{\gamma}$ based on (7.46), the subspace minimization method [123] is employed due to its computational efficiency, which is equivalent to the joint estimation of all the unknowns. Let us define $\mathbf{D} = [[\mathbf{B}]_{:,1:M-1},\ \mathbf{C}]$ of size $M^2 \times (2M - 1)$, and obtain an orthogonal projection matrix onto the orthogonal complement of $\mathbf{D}$ as $\mathbf{P}_d = \mathbf{I}_{M^2} - \mathbf{D}(\mathbf{D}^T\mathbf{D})^{-1}\mathbf{D}^T$, which fullfills the condition that

$\mathbf{P}_d\mathbf{B} = \mathbf{P}_d\mathbf{C} = \mathbf{0}_{M^2 \times M}$. Premultiplying $\mathbf{P}_d$ to both sides of (7.46), we arrive at

$$\mathbf{P}_d\mathbf{A}\boldsymbol{\gamma} = -\mathbf{P}_d\mathbf{g}. \tag{7.47}$$

Consequently, the LS estimate of $\boldsymbol{\gamma}$ is given by

$$\hat{\boldsymbol{\gamma}} = -(\mathbf{A}^T\mathbf{P}_d\mathbf{A})^{-1}\mathbf{A}^T\mathbf{P}_d\mathbf{g}. \tag{7.48}$$

We remark that in order to obtain the LS estimate of $\boldsymbol{\gamma}$, the condition $M^2 + 1 - 2M \geq M$ has to be fullfilled taking the penalty of the projection into account, which means $M_{\min} = 3$ is still valid.

In the second step, plugging $\hat{\boldsymbol{\gamma}}$ into (7.46), we achieve

$$\mathbf{A}\hat{\boldsymbol{\gamma}} = (\mathbf{B} + \mathbf{C})\mathbf{d} + \mathbf{C}\boldsymbol{\Delta} - \mathbf{g}, \tag{7.49}$$

where we would like to get rid of the nuisance parameter $\boldsymbol{\Delta}$, and investigate the nonlinear relation between $\mathbf{d}$ and $\mathbf{x}$ in order to obtain a unique estimate of $\mathbf{x}$. Thus, the orthogonal projection matrix $\mathbf{P}_c$ onto the orthogonal complement of $\mathbf{C}$ is used, which is given by $\mathbf{P}_c = \mathbf{I}_{M^2} - \frac{1}{M}(\mathbf{I}_M \otimes (\mathbf{1}_M\mathbf{1}_M^T))$. Moreover, we find that $\mathbf{P}_c\mathbf{B}\mathbf{d} = \mathbf{B}\mathbf{d} - \bar{d}\mathbf{1}_{M^2}$, where $\bar{d} = \frac{1}{M}\sum_{i=1}^{M} d_i$. Consequently, premultiplying $\mathbf{P}_c$ to both sides of (7.49), we obtain

$$\mathbf{P}_c\mathbf{A}\hat{\boldsymbol{\gamma}} = \mathbf{B}\mathbf{d} - \bar{d}\mathbf{1}_{M^2} - \mathbf{P}_c\mathbf{g}. \tag{7.50}$$

Due to the special structure of $\mathbf{B}$, we have $\frac{1}{M}\mathbf{B}^T\mathbf{B} = \mathbf{I}_M$. Therefore, premultiplying $\frac{1}{M}\mathbf{B}^T$ to both sides of (7.50), moving $\mathbf{d}$ to one side, the other terms to the other side, and simplifying the equations, we arrive at

$$\mathbf{d} = \frac{1}{M}\mathbf{B}^T\mathbf{P}_c\mathbf{A}\hat{\boldsymbol{\gamma}} + \bar{d}\mathbf{1}_M + \frac{1}{M}\mathbf{B}^T\mathbf{P}_c\mathbf{g}. \tag{7.51}$$

After element-wise multiplication on both sides of the equation, moving unknown parameters to one side, known terms to the other side, we achieve

$$\boldsymbol{\phi}_s = \mathbf{H}_s\mathbf{s}_s, \tag{7.52}$$

where $\boldsymbol{\phi}_s = \boldsymbol{\psi}_a - \frac{1}{M^2}(\mathbf{B}^T\mathbf{P}_c(\mathbf{A}\hat{\boldsymbol{\gamma}} + \mathbf{g})) \odot (\mathbf{B}^T\mathbf{P}_c(\mathbf{A}\hat{\boldsymbol{\gamma}} + \mathbf{g}))$,
$\mathbf{H}_s = [2\mathbf{X}_a^T, \ \mathbf{1}_M, \ \frac{2}{M}\mathbf{B}^T\mathbf{P}_c(\mathbf{A}\hat{\boldsymbol{\gamma}} + \mathbf{g})]$ and $\mathbf{s}_s = [\mathbf{x}^T, \ \bar{d}^2 - \|\mathbf{x}\|^2, \ \bar{d}]^T$. Thus, the LS estimate of $\mathbf{s}_s$ is given by

$$\hat{\mathbf{s}}_s = (\mathbf{H}_s^T\mathbf{H}_s)^{-1}\mathbf{H}_s^T\boldsymbol{\phi}_s. \tag{7.53}$$

We remark that the rank of $\mathbf{H}_s$ should be $l + 2$, and thus $M \geq l + 2$, e.g., $M \geq 4$ for $l = 2$. One more anchor is required due to the linearization compared to $M_{\min} = 3$ mentioned before.
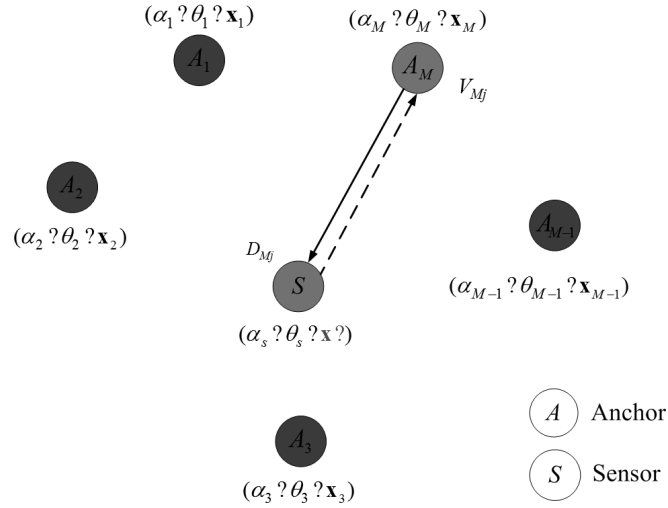
Figure 7.4: An illustration of the TWR protocol

7.5 Localization using the TWR protocol

7.5.1 System Model

Since the TWR protocol is proposed in the standard [75], we also develop an approach based on this protocol to jointly synchronize and localize fully asynchronous networks thereby overcoming some of the drawbacks in [149,151]. In [149,151], the consecutive [149] or periodic [151] transmission of ranging packets is required in order to make use of the prior knowledge of the packet length for relative clock skew estimation. Moreover, due to the fact that the relative clock skew is expressed in ppm (10^{-6}), [149, 151] ignores the clock drift during the time of flight, which is in the order of tens of nanoseconds for an indoor environment, and thus [149, 151] introduces an approximation. In our proposed scheme, we do not require a consecutive or periodic packet transmission, and do not ignore any clock drift. Furthermore, closed-form LS estimators are developed with computational efficiency.

Each anchor carries out multiple iterations of the TWR protocol to measure its distance to the target node as shown in Fig. 7.4. As a result, the ith anchor measures its round trip time at the jth iteration of the TWR protocol as V_{ij}, which is obtained by making the difference of its two timestamps recorded upon the departure and the arrival of the RMARKERs of the ranging request and the ranging response, respectively. The target node correspondingly measures its processing time D_{ij} in the same way. The clock offset of the node is eliminated in the same way as for the ATR protocol by making the difference of its timestamps.

However, the clock skews of the nodes still remain. Thus the relation between V_{ij} and D_{ij} for the jth iteration can be modeled as

$$\frac{c}{2}\left(\frac{V_{ij}}{\alpha_i} - \frac{D_{ij}}{\alpha_s}\right) = d_i + \frac{\breve{n}_{ij}}{\alpha_i} - \frac{\breve{m}_{ij}}{\alpha_s}, \; i = 1, 2, \ldots, M, \qquad (7.54)$$

where $\breve{n}_{ij}$ and $\breve{m}_{ij}$ are the distance error terms translated from the measurement errors in V_{ij} and D_{ij}, which can be modeled as zero-mean Gaussian random variables. Once all the anchors execute the TWR protocol for one time, we obtain M equations, but $M + l + 1$ unknowns in total. Hence, the TWR protocol is executed N times by each anchor to obtain extra information. Defining $\mathbf{p}_i = \frac{c}{2}[V_{i1}, \; V_{i2}, \; \ldots, \; V_{iN}]^T$ and $\mathbf{q}_i = \frac{c}{2}[D_{i1}, \; D_{i2}, \; \ldots, \; D_{iN}]^T$ obtained by the ith anchor-target pair, we arrive at

$$\frac{1}{\alpha_i}\mathbf{p}_i - \frac{1}{\alpha_s}\mathbf{q}_i = d_i\mathbf{1}_N + \frac{1}{\alpha_i}\breve{\mathbf{n}}_i - \frac{1}{\alpha_s}\breve{\mathbf{m}}_i, \; i = 1, 2, \ldots, M. \qquad (7.55)$$

Note that for quasi-synchronous networks, where $\alpha_i = 1, i = 1, 2, \ldots, M$, if we have $\alpha_s = 1$, (7.55) boils down to the following equation after averaging

$$\frac{1}{N}\mathbf{1}_N^T(\mathbf{p}_i - \mathbf{q}_i) = d_i + \frac{1}{N}\mathbf{1}_N^T(\breve{\mathbf{n}}_i - \breve{\mathbf{m}}_i), \; i = 1, 2, \ldots, M. \qquad (7.56)$$

We can write (7.56) in vector form collecting all the timestamps for each anchor-target pair. Consequently, the target position can then be estimated based on the vector form of (7.56) using the same LS estimator as in Section 7.3. We do not repeat it here for the interest of brevity. We remark that we first have to calibrate the clock skew of the target clock in a quasi-synchronous network in order to employ the TWR protocol. Otherwise, the target clock skew will influence its processing time record and corrupt the location estimate.

7.5.2   Localization algorithm

Now let us explore closed-form localization methods for fully asynchronous networks using the TWR protocol. For simplicity, we ignore the noise terms from now on. The equations (7.55) for different anchor-target pairs are coupled through α_s and $\mathbf{x}$. They are nonlinear w.r.t. $\mathbf{x}$ due to the nonlinear relation $\mathbf{d} \odot \mathbf{d} = \psi_a - 2\mathbf{X}_a^T\mathbf{x} + \|\mathbf{x}\|^2\mathbf{1}_M$, but each of them is linear w.r.t. $1/\alpha_i$, $1/\alpha_s$ and d_i. Therefore, we can first estimate the relative clock skew $\beta_{si} = \alpha_s/\alpha_i$ based on a set of equations modified from (7.55):

$$\beta_{si}\mathbf{p}_i - \mathbf{q}_i = \alpha_s d_i\mathbf{1}_N. \qquad (7.57)$$

Note that α_s and d_i are coupled together. According to (7.57), $N \geq 2$ in order to estimate both β_{si} and $\alpha_s d_i$. Since we are only interested in β_i, an orthogonal projection matrix $\mathbf{P}_N = \mathbf{I}_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^T$ onto the complement of $\mathbf{1}_N$ is constructed.

Applying $\mathbf{P}_N$ to both sides of (7.57), we can estimate β_{si} as

$$\hat{\beta}_{si} = \frac{\mathbf{p}_i^T \mathbf{P}_N \mathbf{q}_i}{\mathbf{p}_i^T \mathbf{P}_N \mathbf{p}_i}. \tag{7.58}$$

We remark that a different processing time is again required in each iteration of the TWR protocol. Otherwise, $\mathbf{p}_i$ and $\mathbf{q}_i$ would be canceled out by the projection in the noiseless case. Sequentially, we plug $\hat{\beta}_{si}$ into (7.57) and average it to mitigate the noise, which leads to

$$\frac{1}{N\alpha_s} \mathbf{1}_N^T (\hat{\beta}_{si} \mathbf{p}_i - \mathbf{q}_i) = d_i, \ i = 1, 2, \ldots, M. \tag{7.59}$$

After element-wise multiplication of (7.59), moving knowns to one side and unknowns to the other side, we achieve

$$\boldsymbol{\psi}_a = \mathbf{H}_t \mathbf{s}_t, \tag{7.60}$$

where $\mathbf{s}_t = [\mathbf{x}^T, \ \|\mathbf{x}\|^2, \ \frac{1}{\alpha_s^2}]^T$, $\mathbf{H}_t = [2\mathbf{X}_a^T, \ -\mathbf{1}_M, \ \mathbf{f}]$ and the ith element of $\mathbf{f}$ is defined as $[\mathbf{f}]_i = \frac{1}{N^2}(\mathbf{1}_N^T(\hat{\beta}_{si}\mathbf{p}_i - \mathbf{q}_i)) \odot (\mathbf{1}_N^T(\hat{\beta}_{si}\mathbf{p}_i - \mathbf{q}_i))$. Consequently, the LS estimate of $\mathbf{s}_t$ is given by

$$\hat{\mathbf{s}}_t = (\mathbf{H}_t^T \mathbf{H}_t)^{-1} \mathbf{H}_t^T \boldsymbol{\psi}_a. \tag{7.61}$$

We remark that the rank of $\mathbf{H}_t$ should be $l + 2$, thus $M \geq l + 2$, e.g., $M \geq 4$ for $l = 2$.

Let us now compare the communication load of this approach with localization approach II using the ATR protocol in Section 7.4.2. Note that localization approach I using the ATR protocol has an even lower communication load than localization approach II. Thus, we just compare this approach with localization approach II. In the worst case, the ATR protocol is executed by every anchor in the network (in total M anchors) in localization approach II. Therefore, $2M$ ranging packets are transmitted, and $2M^2$ timestamps are recorded. On the other hand, if we run the minimum number of iterations for each anchor-target pair in the TWR protocol, which is $N_{\min} = 2$, $2N_{\min}M$ ranging packets are transmitted, and $4N_{\min}M$ timestamps are recorded. Obviously, using localization approach II based on the ATR protocol, we obtain more information and have a smaller communication load. Moreover, the estimate of $\boldsymbol{\gamma}$ is based on the whole set of measurements, but the estimate of β_{si} only depends on a subset of measurements. Furthermore, the computational complexities of the estimator (7.53) for $\mathbf{s}_s$ and (7.61) for $\mathbf{s}_t$ are similar, while the one of the estimator (7.58) for β_{si} is smaller than the one of (7.48) for $\boldsymbol{\gamma}$.

## 7.6   Performance bounds and simulation results

As a well-adopted lower bound, the Cramèr-Rao bound (CRB) is derived for quasi-synchronous and fully asynchronous networks, respectively. Although the CRB is a bound for unbiased estimators, the estimators derived in this chapter are biased. It is still interesting to compare them. Here, we exemplify the CRBs for location estimation on a plane, e.g., we take $l = 2$. The Fisher information matrix (FIM) $\mathbf{I}_1(\boldsymbol{\theta}_1)$ based on the model (7.5) in Section 7.3.1 for quasi-synchronous networks is derived in Appendix 7.C, where $\boldsymbol{\theta}_1 = [\Delta_M, \mathbf{x}^T]^T$, and $\mathbf{x} = [x_1, x_2]^T$. Consequently, we obtain $\text{CRB}(x_1) = [\mathbf{I}_1^{-1}(\boldsymbol{\theta}_1)]_{2,2}$ and $\text{CRB}(x_2) = [\mathbf{I}_1^{-1}(\boldsymbol{\theta}_1)]_{3,3}$. We observe that Δ_M is not part of $\mathbf{I}_1^{-1}(\boldsymbol{\theta}_1)$. Therefore, no matter how large Δ_M is, it has the same influence on the CRB for quasi-synchronous networks. On the other hand, the FIM $\mathbf{I}_2(\boldsymbol{\theta}_2)$ based on the model (7.25)-(7.27) in Section 7.4.1 for fully asynchronous networks employing localization approach I is derived in Appendix 7.D, where $\boldsymbol{\theta}_2 = [\mathbf{x}^T, \Delta_M, \boldsymbol{\alpha}^T, 1/\alpha_s]^T$, and $\boldsymbol{\alpha} = [\alpha_1, \alpha_2, \ldots, \alpha_M]^T$. As a result, we achieve $\text{CRB}(x_1) = [\mathbf{I}_2^{-1}(\boldsymbol{\theta}_2)]_{1,1}$ and $\text{CRB}(x_2) = [\mathbf{I}_2^{-1}(\boldsymbol{\theta}_2)]_{2,2}$. All the parameters of $\boldsymbol{\theta}_2$ appear in $\mathbf{I}_2(\boldsymbol{\theta}_2)$, and thus they all influence the CRB for fully asynchronous networks.

Let us now evaluate the performance of the proposed robust localization algorithms by Monte Carlo simulations, and compare it with the CRB. We consider two simulation setups: Setup 1 and Setup 2. In Setup 1, the anchors are evenly located on the edges of a $40\,\text{m} \times 40\,\text{m}$ rectangular to mimic an indoor geometry scale. Meanwhile the target node is randomly located on a grid with cells of size $1\,\text{m} \times 1\,\text{m}$ inside the rectangular. In Setup 2, all anchors and the target node are randomly distributed on the grid inside the rectangular. Furthermore, m_{ij}, m_{ji} and $\widetilde{m}_{ij}, i, j \in \{1, 2, \ldots, M\}$ have the same variance $\sigma_{ij}^2 = \sigma_{ji}^2$, while $n_{ij}, \widetilde{n}_{ij}$, $\breve{n}_{il}$ and $\breve{m}_{il}, l \in \{1, 2, \ldots, N\}$ have the same variance σ_i^2. Due to the broadcast property of the ranging protocol, we assume that σ_i^2 and σ_{ij}^2 are related to the distances according to the path loss law. Thus we define the average noise power as $\bar{\sigma}^2 = 1/M \sum_{i=1}^{M} \sigma_i^2$, where σ_{ij}^2 and σ_i^2 are chosen to fulfill the condition that all σ_{ij}^2/d_{ij}^2 and σ_i^2/d_i^2 are equal as in [11]. Note that since $d_{ii} = 0$, we simply assume $\sigma_{ii}^2 = 0$ and $m_{ii} = \widetilde{m}_{ii} = 0$. The clock skews of the anchors and the target are randomly generated in the range of $[1 - 100\,\text{ppm}, 1 + 100\,\text{ppm}]$, and the clock offsets are randomly generated in the range of [1ns, 10ns]. For localization approach I, the processing time of the target node in response to the Mth anchor is 5 ms, and as a result the corresponding distance Δ_M is $3 \times 10^8 \times 5 \times 10^{-3} = 1.5 \times 10^6$ m. The time interval T_I between RMARKER1 and RMARKER2 is 1 ms. For localization approach II, the processing time at the target node is randomly generated, uniformly distributed in the range of 2.5 ms to 7.5 ms. As a result, the corresponding distance Δ_i is in the range of 7.5×10^5 m ($3 \times 10^8 \times 2.5 \times 10^{-3}$) to 2.25×10^6 m. The performance criterion is the root mean square error (RMSE) of $\hat{\mathbf{x}}$ vs. SNR, which can be expressed as $\sqrt{1/N_{exp} \sum_{j=1}^{N_{exp}} \|\hat{\mathbf{x}}^{(j)} - \mathbf{x}\|^2}$, where $\hat{\mathbf{x}}^{(j)}$ is the estimate obtained in the jth trial. Each simulation result is averaged

over $N_{exp} = 1000$ Monte Carlo trials. We would like to compare our localization algorithms with the localization algorithm using the TWR protocol, which is clarified in Section 7.5, and the algorithm LS-I in [28]. We assume that the algorithm LS-I in [28] is employed with accurate knowledge of the anchor clock parameters and positions. The number of rounds of two-way message exchanges for LS-I in [28] is four, as with more than four rounds, the estimation performance improvement is only marginal [28].

7.6.1   Localization for quasi-synchronous networks

Fig. 7.5(a) and Fig. 7.5(b) show the localization performance of respectively Setup 1 and Setup 2 with eight anchors that do not suffer from clock skews. In each Monte Carlo run, we generate a new geometry. In both figures, the dashed lines with no and "$\diamondsuit$" markers represent the conventional localization algorithm with the TWR protocol using the fraudulent timestamp report from the target node with 3 m and 15 m errors, respectively. According to the figures, they cannot estimate the target position correctly even with sufficiently small noise terms. A larger timestamp error introduces a higher error floor. The dashed line with "$\triangledown$" markers illuminates the conventional localization algorithm with the TWR protocol using the correct timestamp report. It is slightly better than the CRB of our method for Setup 1 (the dotted line with "$\times$" markers), but much better than the one for Setup 2. This is reasonable, since the conventional method estimates less parameters than the proposed method. The performance of the algorithm LS-I (the dashed line with "$\triangle$" markers) in Setup 1 and Setup 2 is quite different. It is worse than our method (the solid lines with different markers) in Setup 1, whereas it is better than our method in Setup 2. As a result, the algorithm LS-I seems to be sensitive to the geometry, when the target node is inside the region restricted by the anchors. Moreover, $8M$ ranging packets are transmitted in the algorithm LS-I compared to only 2 ranging packets in our scheme, so our communication load is much smaller. Furthermore, the method in Section 7.3.2 is immune to a fraudulent timestamp report and robust to the randomness of the target node clock. Moreover, its localization performance is accurate with sufficiently small noise terms. The solid line with "$\square$" markers shows the performance of the LS estimator using the eighth anchor as the reference node [135], whereas the solid line with "$+$" markers indicates the performance of our proposed LS estimator using the projection $\mathbf{P}$. Note that they almost overlap. The solid lines with "$\circ$" and "$*$" markers denote the performance of our proposed WLS method with an optimal weighting and an iterative weighting matrix, respectively. The fact that they almost overlap indicates that if we use the LS estimate as an initial point, the iterative WLS can converge to the WLS with optimal weighting. The performance of the WLS with optimal weighting is slightly better than the LS and the iterative WLS estimators. Considering the computational complexity and the performance, the LS estimator would be the best option.
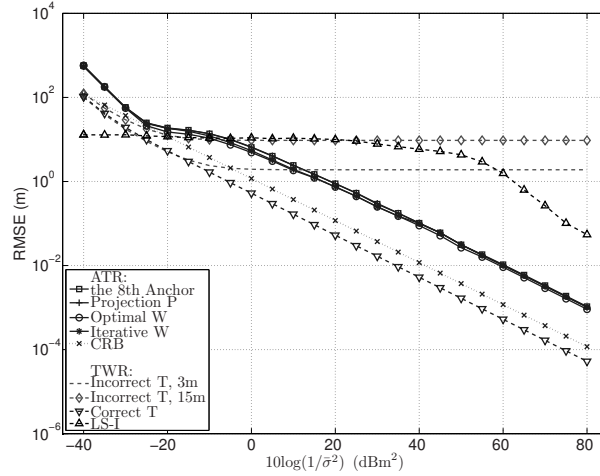
(a) Setup 1, $M = 8$
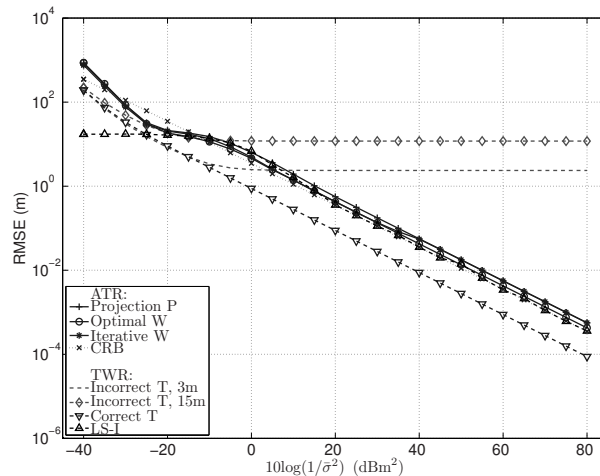
(b) Setup 2, $M = 8$

Figure 7.5: RMSE of target node position $\mathbf{x}$ for quasi-synchronous networks

7.6.2   Localization for fully asynchronous networks

Localization approach I

Fig. 7.6(a) and Fig. 7.6(b) illustrate the localization performance of respectively Setup 1 and Setup 2 with eight anchors in fully asynchronous networks using localization approach I. The dotted lines with "$\diamondsuit$" markers depict the performance

of the conventional TWR algorithm without clock skew calibration. The high error floors indicate that it cannot estimate the correct position of the target even with sufficiently small noise. The dotted line with "$\triangledown$" markers indicates the performance of the conventional TWR algorithm using a correct timestamp report and with clock skew calibration. It is lower than the dotted line with "$\times$" markers, which represents the CRB and serves as a benchmark for fully asynchronous networks. There are performance gaps between the CRB and the proposed methods in both figures. However, the gap is smaller in Setup 2. The proposed methods make a tradeoff between performance and complexity. The performance of the algorithm LS-I (the dashed line with "$\triangle$" markers) in Setup 1 and Setup 2 is again quite different. It is slightly worse than our method (the solid lines with different markers) in Setup 1, whereas it is better than the CRB in Setup 2. Note that it estimates much fewer parameters than ours, since it assumes the exact knowledge of the anchor clock parameters. On the other hand, we estimate the anchor clock skews and the target node position together. The algorithm LS-I still seems to be sensitive to the geometry, and has much more communication load than ours. The estimator $\hat{\mathbf{s}}_a$ (7.38) (the lines with "$+$" markers) and the CCS-ENP method (the lines with "$\square$" markers) achieve the same performance. The performance of the estimator $\hat{\mathbf{s}}_d$ (7.35) (the lines with "$\circ$" markers) is worse than them in general, as it estimates more parameters. Note that $\hat{\mathbf{s}}_a$ treats α_s as 1, and $\hat{\mathbf{s}}_d$ treats $1/\alpha_s^2$ as $2/\alpha_s - 1$ and $\bar{d}/\alpha_s$ as $\bar{d}$. Therefore, there are error floors when $\alpha_s \neq 1$ (the solid lines with "$+$" and "$\circ$" markers), but these floors only appear at high SNR. The performance degradation at high SNR caused by the approximation error in $\hat{\mathbf{s}}_a$ is more important than the benefit of less unknowns. Thus, $\hat{\mathbf{s}}_a$ has a slightly higher error floor than $\hat{\mathbf{s}}_d$ when $\alpha_s \neq 1$ at high SNR. However, as long as the condition $\alpha_s = 1$ is fulfilled, both $\hat{\mathbf{s}}_a$ and $\hat{\mathbf{s}}_d$ (the dashed lines with "$+$" and "$\circ$" markers) can achieve accurate estimates, even for high SNR. The CCS-ENP method is developed by treating α_M as 1. Thus, it has a similar error floor at high SNR when $\alpha_M \neq 1$, which is indicated by the solid line with "$\square$" markers. Based on the above analysis, both the estimator $\hat{\mathbf{s}}_a$ and the CCS-ENP method are good choices considering the implementation cost and performance.

Localization approach II

We now compare localization approach II with the localization method using the TWR protocol proposed in Section 7.5. We only consider Setup II, where all the anchors and the target node are randomly distributed inside a $40\,\text{m} \times 40\,\text{m}$ rectangular. Fig. 7.7(a) and Fig. 7.7(b) show the RMSE of $\mathbf{x}$ vs. $1/\bar{\sigma}^2$ for both protocols, respectively. We have tested different numbers of anchors. As $M = 5$ is just one more than the minimum number of anchors required by the approaches, the curves with $M = 5$ (the lines with circle markers) are not as smooth as the ones with more anchors. More anchors improve the accuracy of the estimates. In Fig. 7.7(a), the performance gap between $M = 5$ and $M = 7$ (the line with
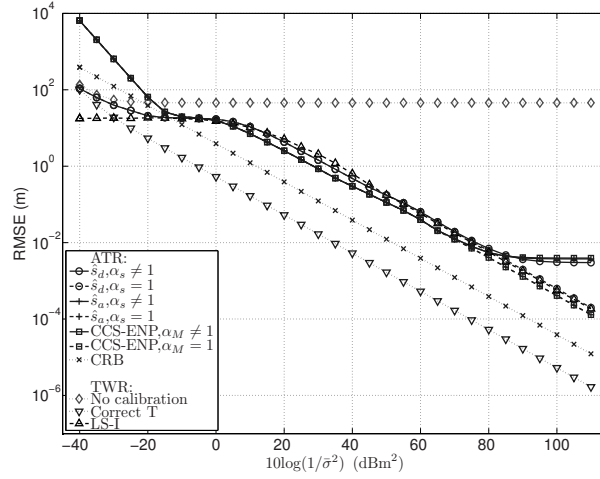
(a) Setup 1, $M = 8$
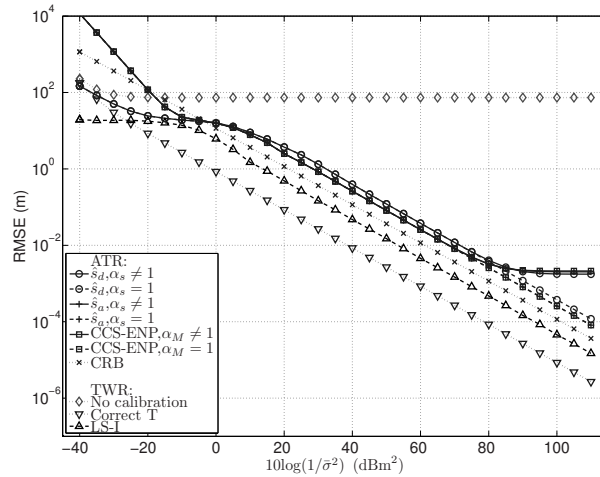
(b) Setup 2, $M = 8$

Figure 7.6: RMSE of target node position **x** for fully asynchronous networks

+ markers) is larger than the one between $M = 7$ and $M = 9$ (the line with rectangular markers). Thus the improvement reduces as the number of anchors increases. In Fig. 7.7(b), we choose $N = 3$. We observe that the improvement first increases and then reduces along with the number of anchors. Although the approach based on the TWR protocol transmits $N - 1$ times more packets than the one based on the ATR protocol, we observe in general better performance
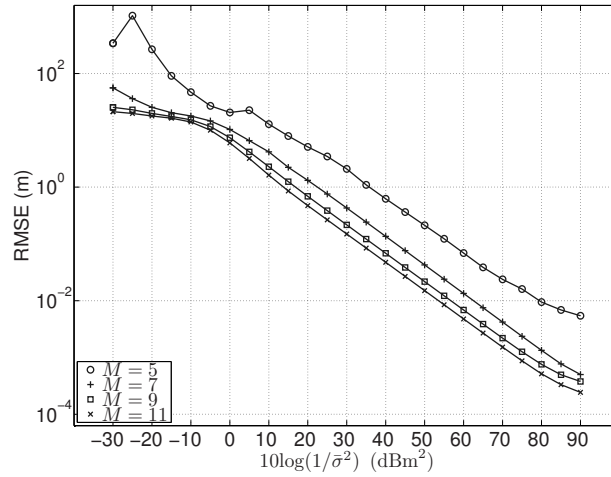
(a) The ATR protocol
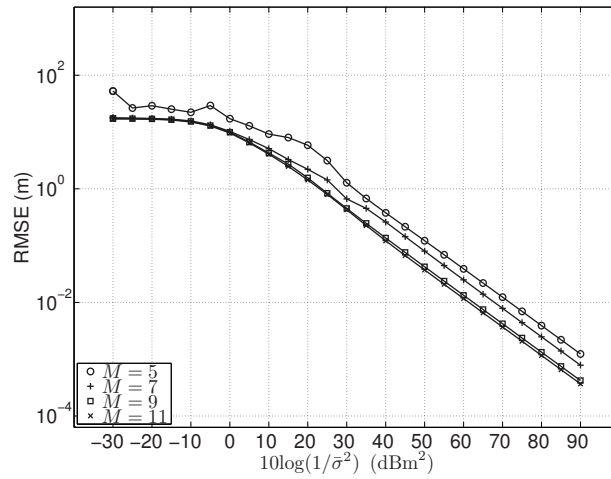
(b) The TWR protocol

Figure 7.7: RMSE of **x** for different protocols

achieved by the latter one.

7.7 Conclusions

In this chapter, we have proposed robust localization strategies based on TOA measurements to localize a target node with the help of anchors for asynchronous

	ATR	TWR
Broadcast property	Yes	No
Internal attack by the target node	Immune	Vulnerable
Processing Time of the target node	Unknown	Measured
Communication load	Low	High

Table 7.1: Comparison between the ATR and the TWR protocol

networks. We have dealt with two kinds of asynchronous networks: one only with clock offsets referred to as quasi-synchronous networks, and the other not only with clock offsets but also with clock skews referred to as fully asynchronous networks. Regardless of the reliability of the timestamps from the target node, we have proposed a novel ranging protocol, namely asymmetric trip ranging (ATR), which leads to localization methods that are naturally immune to internal attacks mounted by a compromised target node. The comparison between the ATR and the TWR protocol is listed in Table 7.1. Furthermore, closed-form LS and iterative WLS estimators have been proposed to localize the target node for quasi-synchronous networks. For fully asynchronous networks, we have developed two different approaches to jointly estimate the position and clock parameters. Localization approach I with the minimum communication load takes practical issues into account to simplify the estimators by appropriate approximations. The price paid by localization approach I is the error floor at high SNR due to approximations. Localization approach II relaxes the communication load constraint in order to obtain accurate estimates and avoid the error floor at high SNR. The appropriate estimators can be chosen according to system specifications. We have further designed closed-form LS estimators for joint synchronization and localization using the TWR protocol. Moreover, the CRBs for both quasi-synchronous and fully asynchronous networks have been derived, respectively. Simulation results have corroborated the efficiency of our localization methods.

7.A Computation of $\mathbf{K}^{-1}$

Since $\mathbf{K}_1 = \mathbf{I}_M - \mathbf{F}_3(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}_3^T$, $\mathbf{K}_1^{-1}$ can be written as

$$
\begin{aligned}
\mathbf{K}_1^{-1} &= (\mathbf{I}_M - \mathbf{F}_3(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}_3^T)^{-1} \\
&= (\mathbf{I}_M - ((\mathbf{F}_3^T)^{-1}(\mathbf{F}^T\mathbf{F})\mathbf{F}_3^{-1})^{-1})^{-1} \\
&= (\mathbf{F}_3^T)^{-1}(\mathbf{F}^T\mathbf{F})\mathbf{F}_3^{-1}((\mathbf{F}_3^T)^{-1}(\mathbf{F}^T\mathbf{F})\mathbf{F}_3^{-1} - \mathbf{I}_M)^{-1} \\
&= (\mathbf{F}_3^T)^{-1}(\mathbf{F}^T\mathbf{F})((\mathbf{F}_3^T)^{-1}(\mathbf{F}^T\mathbf{F}) - \mathbf{F}_3)^{-1} \\
&= (\mathbf{F}_3^T)^{-1}(\mathbf{F}^T\mathbf{F})(\mathbf{F}^T\mathbf{F} - \mathbf{F}_3^T\mathbf{F}_3)^{-1}\mathbf{F}_3^T.
\end{aligned}
\tag{7.62}
$$

Recalling that $\mathbf{F} = [\mathbf{F}_1^T\mathbf{Z}^T - \mathbf{E}_1^T,\ \mathbf{F}_2^T,\ \mathbf{F}_3^T]^T$, we arrive at

$$
\begin{aligned}
\mathbf{F}^T\mathbf{F} - \mathbf{F}_3^T\mathbf{F}_3 &= \mathbf{F}_1^T\mathbf{Z}^T\mathbf{Z}\mathbf{F}_1 + \mathbf{F}_2^T\mathbf{F}_2 + \mathbf{E}_1^T\mathbf{E}_1 - \mathbf{F}_1^T\mathbf{Z}^T\mathbf{E}_1 - \mathbf{E}_1^T\mathbf{Z}\mathbf{F}_1 \\
&= \begin{bmatrix} \mathbf{\Lambda} & \boldsymbol{\rho} \\ \boldsymbol{\rho}^T & f \end{bmatrix},
\end{aligned}
\tag{7.63}
$$

where $\boldsymbol{\rho} = cT_I\mathbf{Z}(\widetilde{\mathbf{v}}_M - \mathbf{v}_M)$, $f = (M-1)c^2T_I^2 + [\mathbf{r}_2]_M$, $\mathbf{\Lambda} = \mathbf{Z}\big(\mathrm{diag}(\mathbf{r}_1 + \mathbf{r}_2)\big)\mathbf{Z}^T$, which is a diagonal matrix of size $(3M-1) \times (3M-1)$ with $\mathbf{r}_1 = (\widetilde{\mathbf{v}}_M - \mathbf{v}_M) \odot (\widetilde{\mathbf{v}}_M - \mathbf{v}_M)$ and $\mathbf{r}_2 = (\widetilde{\mathbf{u}}_M - \mathbf{u}_M) \odot (\widetilde{\mathbf{u}}_M - \mathbf{u}_M)$. The inverse of $\mathbf{F}^T\mathbf{F} - \mathbf{F}_3^T\mathbf{F}_3$ is given by

$$
(\mathbf{F}^T\mathbf{F} - \mathbf{F}_3^T\mathbf{F}_3)^{-1} = \begin{bmatrix} \mathbf{\Lambda}^{-1} + \dfrac{1}{w}\mathbf{\Lambda}^{-1}\boldsymbol{\rho}\boldsymbol{\rho}^T\mathbf{\Lambda}^{-1} & \dfrac{1}{w}\mathbf{\Lambda}^{-1}\boldsymbol{\rho} \\ \dfrac{1}{w}\boldsymbol{\rho}^T\mathbf{\Lambda}^{-1} & \dfrac{1}{w} \end{bmatrix}, \tag{7.64}
$$

where $w = f - \boldsymbol{\rho}^T\mathbf{\Lambda}^{-1}\boldsymbol{\rho}$. We further have

$$
\mathbf{\Lambda}^{-1} = \mathbf{Z}\big(\mathrm{diag}(\mathbf{r}_1 + \mathbf{r}_2)\big)^{-1}\mathbf{Z}^T,
$$

$$
\mathbf{\Lambda}^{-1}\boldsymbol{\rho} = cT_I\mathbf{Z}\big(\mathrm{diag}(\mathbf{r}_1 + \mathbf{r}_2)\big)^{-1}\mathbf{Z}^T\mathbf{Z}(\widetilde{\mathbf{v}}_M - \mathbf{v}_M),
$$

$$
\begin{aligned}
\mathbf{\Lambda}^{-1}\boldsymbol{\rho}\boldsymbol{\rho}^T\mathbf{\Lambda}^{-1} = {}& c^2T_I^2\mathbf{Z}\big(\mathrm{diag}(\mathbf{r}_1 + \mathbf{r}_2)\big)^{-1}\mathbf{Z}^T\mathbf{Z}(\widetilde{\mathbf{v}}_M - \mathbf{v}_M) \\
& \times (\widetilde{\mathbf{v}}_M - \mathbf{v}_M)^T\mathbf{Z}^T\mathbf{Z}\big(\mathrm{diag}(\mathbf{r}_1 + \mathbf{r}_2)\big)^{-1}\mathbf{Z}^T,
\end{aligned}
$$

$$
\boldsymbol{\rho}^T\mathbf{\Lambda}^{-1}\boldsymbol{\rho} = c^2T_I^2(\widetilde{\mathbf{v}}_M - \mathbf{v}_M)^T\mathbf{Z}^T\mathbf{Z}\big(\mathrm{diag}(\mathbf{r}_1 + \mathbf{r}_2)\big)^{-1}\mathbf{Z}^T\mathbf{Z}(\widetilde{\mathbf{v}}_M - \mathbf{v}_M).
$$

$\mathbf{K}_1^{-1}$ can be obtained by plugging the expression of $(\mathbf{F}^T\mathbf{F} - \mathbf{F}_3^T\mathbf{F}_3)^{-1}$ into (7.62).

7.B   Computation of H

In the noiseless case or with sufficiently small noise, $\mathbf{Ph} = \alpha_s(\mathbf{Pd} - \mathbf{Pg}_M)$. Recalling that $\mathbf{H} = [2\mathbf{X}_a^T, 2\mathbf{Pg}_M, 2(\mathbf{Ph}) \odot (\mathbf{Pg}_M), \mathbf{1}_M, (\mathbf{Ph}) \odot (\mathbf{Ph}), \mathbf{Ph}]$, we explore the properties of $\mathbf{H}$. With $\mathbf{Pd} = \mathbf{d} - \bar{d}\mathbf{1}_M$, $\mathbf{Pg}_M = \mathbf{g}_M - \bar{d}_a\mathbf{1}_M$ (where $\bar{d}_a = \frac{1}{M}\sum_{i=1}^M d_{iM}$), $\mathbf{d} \odot \mathbf{d} = \boldsymbol{\psi}_a - 2\mathbf{X}_a^T\mathbf{x} + \|\mathbf{x}\|^2\mathbf{1}_M$, and $\mathbf{g}_M \odot \mathbf{g}_M = \boldsymbol{\psi}_a - 2\mathbf{X}_a^T\mathbf{x}_M + \|\mathbf{x}_M\|^2\mathbf{1}_M$, we obtain

$$
\begin{align}
\mathbf{Pd} \odot \mathbf{Pg}_M &= \mathbf{g}_M \odot \mathbf{d} - \bar{d}_a\mathbf{d} - \bar{d}\mathbf{g}_M + \bar{d}_a\bar{d}\mathbf{1}_M, \tag{7.65}\\
(\mathbf{Ph}) \odot (\mathbf{Pg}_M) &= \alpha_s(\mathbf{Pd} - \mathbf{Pg}_M) \odot (\mathbf{Pg}_M) \tag{7.66}\\
&= \alpha_s\big(\mathbf{g}_M \odot \mathbf{d} - \mathbf{g}_M \odot \mathbf{g}_M - \bar{d}_a\mathbf{d} - (\bar{d} + 2\bar{d}_a)\mathbf{g}_M \\
&\quad + (\bar{d}_a\bar{d} - \bar{d}_a^2)\mathbf{1}_M\big),\\
\mathbf{Ph} \odot \mathbf{Ph} &= \alpha_s^2((\mathbf{Pd}) \odot (\mathbf{Pd}) - 2(\mathbf{Pd}) \odot (\mathbf{Pg}_M) + (\mathbf{Pg}_M) \odot (\mathbf{Pg}_M))\\
&= \alpha_s^2(\mathbf{d} \odot \mathbf{d} + \mathbf{g}_M \odot \mathbf{g}_M - 2\mathbf{g}_M \odot \mathbf{d}\\
&\quad + (\bar{d}_a^2 + \bar{d}^2 - 2\bar{d}_a\bar{d})\mathbf{1}_M + 2(\bar{d} - \bar{d}_a)\mathbf{g}_M + 2(\bar{d}_a - \bar{d})\mathbf{d}).
\end{align}
$$
$$\tag{7.67}$$

Based on (7.66) and (7.67), we have

$$
\begin{align}
&\mathbf{Ph} \odot \mathbf{Ph} + 2\alpha_s(\mathbf{Ph}) \odot (\mathbf{Pg}_M)\\
&= \alpha_s^2(2\mathbf{X}_a^T(\mathbf{x}_M - \mathbf{x}) + (\|\mathbf{x}\|^2 - \|\mathbf{x}_M\|^2 + \bar{d}^2 - \bar{d}_a^2)\mathbf{1}_M - 6\bar{d}_a\mathbf{g}_M - 2\bar{d}\mathbf{d}).
\end{align}
$$
$$\tag{7.68}$$

Consequently, the independent columns of $\mathbf{H}$ are $\mathbf{X}_a^T, \mathbf{1}_M, \mathbf{d}, \mathbf{g}_M, \mathbf{Ph} \odot \mathbf{Pg}_M$. The rank of $\mathbf{H}$ is $l + 4$, but its size is $M \times (l + 5)$. Thus, $\mathbf{H}$ is rank-deficient.

7.C   CRB derivation for quasi-synchronous networks

We analyze the CRB for jointly estimating Δ_M and $\mathbf{x}$ based on (7.5). The FIM $\mathbf{I}_1(\boldsymbol{\theta}_1)$ is employed, with entries defined as:

$$
[\mathbf{I}_1(\boldsymbol{\theta}_1)]_{ij} = -E\left[\frac{\partial^2}{\partial[\boldsymbol{\theta}_1]_i\partial[\boldsymbol{\theta}_1]_j}\ln p(\mathbf{u}_M, \mathbf{v}_M; \boldsymbol{\theta}_1)\right], \tag{7.69}
$$

where

$$
p(\mathbf{u}_M, \mathbf{v}_M; \boldsymbol{\theta}_1) = \frac{\exp\left(-\sum_{i=1}^M \frac{(u_i - v_i - d_i - d_M + d_{iM} - \Delta_M)^2}{2(\sigma_i^2 + \sigma_{iM}^2)}\right)}{\sqrt{(2\pi)^M \prod_{i=1}^M (\sigma_i^2 + \sigma_{iM}^2)}}. \tag{7.70}
$$

In the case of localization on a plane ($l = 2$), $\mathbf{I}_1(\boldsymbol{\theta}_1)$ can be specified as

$$\mathbf{I}_1(\boldsymbol{\theta}_1) \quad = \quad \left[\begin{array}{cc} \mathbf{G} & \mathbf{r} \\ \mathbf{r}^T & k \end{array} \right], \tag{7.71}$$

where $k = \sum_{i=1}^{M} 1/(\sigma_i^2 + \sigma_{iM}^2)$, and

$$\mathbf{G} = \left[\begin{array}{c} \sum_{i=1}^{M} \frac{1}{\sigma_i^2 + \sigma_{iM}^2} \left(\frac{x_1 - x_{1,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_1 - x_{1,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right)^2 \\ \sum_{i=1}^{M} \frac{1}{\sigma_i^2 + \sigma_{iM}^2} \left(\frac{x_2 - x_{2,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_2 - x_{2,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right) \left(\frac{x_1 - x_{1,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_1 - x_{1,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right) \end{array} \right.$$
$$\left. \begin{array}{c} \sum_{i=1}^{M} \frac{1}{\sigma_i^2 + \sigma_{iM}^2} \left(\frac{x_1 - x_{1,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_1 - x_{1,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right) \left(\frac{x_2 - x_{2,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_2 - x_{2,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right) \\ \sum_{i=1}^{M} \frac{1}{\sigma_i^2 + \sigma_{iM}^2} \left(\frac{x_2 - x_{2,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_2 - x_{2,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right)^2 \end{array} \right] \tag{7.72}$$

$$\mathbf{r} = \left[\begin{array}{cc} \sum_{i=1}^{M} \frac{\left(\frac{x_1 - x_{1,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_1 - x_{1,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right)}{\sigma_i^2 + \sigma_{iM}^2} & \sum_{i=1}^{M} \frac{\left(\frac{x_2 - x_{2,i}}{\|\mathbf{x} - \mathbf{x}_i\|} + \frac{x_2 - x_{2,M}}{\|\mathbf{x} - \mathbf{x}_M\|} \right)}{\sigma_i^2 + \sigma_{iM}^2} \end{array} \right]^T.$$
$$\tag{7.73}$$

7.D   CRB derivation for fully asynchronous networks

We rewrite (7.25) - (7.27) as

$$\mathbf{Z}(\widetilde{\mathbf{v}}_M - \mathbf{v}_M) \quad = \quad \frac{cT_I}{\alpha_M} \mathbf{Z}\boldsymbol{\alpha} + \mathbf{Z}(\widetilde{\mathbf{m}}_M - \mathbf{m}_M), \tag{7.74}$$

$$\widetilde{\mathbf{u}}_M - \mathbf{u}_M \quad = \quad \frac{cT_I}{\alpha_s}\boldsymbol{\alpha} + \widetilde{\mathbf{n}}_M - \mathbf{n}_M, \tag{7.75}$$

$$\mathbf{u}_M - \mathbf{v}_M \quad = \quad \boldsymbol{\alpha} \odot (\mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M) + \mathbf{n}_M - \mathbf{m}_M. \tag{7.76}$$

We analyze the CRB for jointly estimating $\mathbf{x}$, Δ_M, $\boldsymbol{\alpha}$ and $1/\alpha_s$ based on (7.74) - (7.76). The FIM $\mathbf{I}_2(\boldsymbol{\theta}_2)$ is employed, with entries defined as:

$$
\begin{aligned}
\mathbf{I}_2(\boldsymbol{\theta}_2) &= -E\left[\frac{\partial^2\ln p(\mathbf{u}_M, \mathbf{v}_M, \widetilde{\mathbf{v}}_M, \widetilde{\mathbf{u}}_M; \boldsymbol{\theta}_2)}{\partial\boldsymbol{\theta}_2\partial\boldsymbol{\theta}_2^T}\right] \\
&= \left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)}{\partial\boldsymbol{\theta}_2}\right]^T \mathbf{C}^{-1} \left[\frac{\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)}{\partial\boldsymbol{\theta}_2}\right],
\end{aligned}
\tag{7.77}
$$

where

$$
\boldsymbol{\mu}(\boldsymbol{\theta}_2) = \begin{bmatrix} (cT_I/\alpha_M)\mathbf{Z}\boldsymbol{\alpha} \\ (cT_I/\alpha_s)\boldsymbol{\alpha} \\ \boldsymbol{\alpha} \odot (\mathbf{d} + (d_M + \Delta_M)\mathbf{1}_M - \mathbf{g}_M) \end{bmatrix},
\tag{7.78}
$$

$$
\mathbf{C} = \begin{bmatrix} 2\mathrm{diag}(\mathbf{Z}\boldsymbol{\sigma}_r) & \mathbf{0}_{(M-1)\times M} & \mathbf{Z}\mathrm{diag}(\boldsymbol{\sigma}_r) \\ \mathbf{0}_{M\times(M-1)} & 2\mathrm{diag}(\boldsymbol{\sigma}_s) & \mathrm{diag}(\boldsymbol{\sigma}_s) \\ \mathrm{diag}(\boldsymbol{\sigma}_r)\mathbf{Z}^T & \mathrm{diag}(\boldsymbol{\sigma}_s) & \mathrm{diag}(\boldsymbol{\sigma}_r + \boldsymbol{\sigma}_s) \end{bmatrix},
\tag{7.79}
$$

where $\boldsymbol{\sigma}_r = [\sigma_{1M}^2, \sigma_{2M}^2, \ldots, \sigma_{(M-1)M}^2, 0]^T$ and $\boldsymbol{\sigma}_s = [\sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2]^T$. Furthermore, we achieve

$$
\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)/\partial\Delta_M = [\mathbf{0}_{M-1}^T, \mathbf{0}_M^T, \boldsymbol{\alpha}^T]^T,
\tag{7.80}
$$

$$
\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)/\partial\frac{1}{\alpha_s} = [\mathbf{0}_{M-1}^T, -cT_I\boldsymbol{\alpha}^T, \mathbf{0}_M^T]^T,
\tag{7.81}
$$

$$
[\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)/\partial\alpha_i]_j = \begin{cases} cT_I/\alpha_M, & j = i \text{ and } i \neq M \\ cT_I/\alpha_s, & j = i + M - 1 \\ d_i + d_M + \Delta_M - d_{iM}, & j = i + 2M - 1 \\ 0 & \text{else} \end{cases},
$$
$$
\text{for } i = 1, \ldots, M, \ j = 1, \ldots, 3M - 1,
\tag{7.82}
$$

$$
[\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)/\partial x_i]_j = \begin{cases} \alpha_{j+1-2M}\left(\dfrac{x_i - x_{i,j+1-2M}}{\|\mathbf{x} - \mathbf{x}_{j+1-2M}\|} + \dfrac{x_i - x_{i,M}}{\|\mathbf{x} - \mathbf{x}_M\|}\right), \\ \qquad\qquad\qquad\qquad\qquad\qquad 2M \leq j \leq 3M - 1, \\ 0, \quad \text{else} \end{cases}
$$
$$
\text{for } i = 1, \ldots, l, \ j = 1, \ldots, 3M - 1.
\tag{7.83}
$$

We obtain $\partial\boldsymbol{\mu}(\boldsymbol{\theta}_2)/\partial\boldsymbol{\theta}_2$ by plugging in the above results. Then we can derive $\mathbf{I}_2(\boldsymbol{\theta}_2)$ based on (7.77).

Tracking a Mobile Node

8.1 Introduction

Tracking a mobile target node is an important issue in many wireless sensor network (WSN) applications [161–163]. In general, tracking systems follow two steps. In the first step, metrics bearing location information are measured, such as time-of-arrival (TOA) or time-difference-of-arrival (TDOA), angle-of-arrival (AOA), and received signal strength (RSS) [1]. High accuracy and potentially low cost implementation make TOA or TDOA based on ultra-wideband impulse radios (UWB-IRs) a promising ranging method [1]. Consequently, clock synchronization has to be taken into account for a localization or a tracking system using TOA or TDOA measurements [136, 137, 162, 163]. In the second step, the ranging measurements are used to track the target position. Due to the nonlinear relations between these ranging measurements and the coordinates of the mobile target node, the conventional Kalman filter (KF) cannot be used. The extended Kalman filter (EKF) [25] is most widely used to linearize the non-linear model. However, the performance of the EKF is decided by how well the linear approximation is. Furthermore, the unscented Kalman filter (UKF) [164] is proposed to overcome the drawbacks of the EKF. The UKF follows the principle that it is easier to approximate a probability distribution than a random nonlinear model, and it calculates the stochastic properties of a random variable undergoing a nonlinear transformation. Moreover, the particle filter [26] is also a powerful tool to deal with nonlinear models and non-Gaussian noise for tracking. However, both the UKF and the particle filter are computationally intensive. An EKF and a UKF

The results in this chapter appeared in [116].

177

are proposed in [162] to track a target node with fixed anchors (nodes with known positions) in asynchronous networks with clock skews and clock offsets. The target node periodically transmits a pulse. The TDOAs of these pulses received by the same anchor are calculated in order to cancel the anchor clock offset. Then the impact of the clock skews on the TDOAs is approximated as a zero mean Gaussian noise term. However, in practice the variation of the clock skew is observed by hours [165], and should thus be viewed as a random variable with an unknown mean rather than a zero mean random variable. A sequential Monte Carlo (SMC) method is proposed in [163] to jointly estimate the clock offsets and the target trajectory for asynchronous WSNs, which is also computationally intensive.

In this chapter, a KF based tracking method is developed to track the target node position with the help of anchors in asynchronous networks with clock offsets. Our work is inspired by [166], where pseudomeasurements linear to the state are constructed based on conventional ranging measurements, and a KF is proposed based on the linear model. But [166] only discusses a scenario, which is composed of three anchor nodes and one target. We consider asynchronous networks with clock offsets among the anchors, and no synchronization requirement for the target node. The asymmetric trip ranging (ATR) protocol proposed in [136, 137] is employed here to obtain TOA measurements and facilitate clock offset cancellation. Since all the TOA measurements are obtained at the anchors, our KF tracker can avoid any influence of the asynchronous target clock. Consequently, a linear measurement model is derived from the TOA measurements via projection and element-wise multiplication. This exact linearization is different from the first order approximation of the EKF. Thus, the KF based on this linear measurement model does not have the modeling errors inherently contained in the EKF. Furthermore, low computational complexity makes the proposed KF a promising solution for practical use. We compare the proposed KF with the EKF by simulations. In future work, we would like to propose low complexity trackers for asynchronous WSNs not only with clock offsets but also with clock skews.

8.2   Linearization of the measurement model

We assume a multiple-anchors-one-target scenario here. We recall that $\mathbf{X}_a$ indicates the known and fixed coordinates of the anchor nodes similarly as the previous chapters. Moreover, a vector $\mathbf{x}(k)$ of length l denotes the unknown coordinates of the target node at time k. In an asynchronous network with clock offsets, the target node clock runs freely, and the clock skews of all the anchors are equal to 1 or treated as 1. There are only clock offsets among all the anchors. Thus, the model for the anchor clock [29] is given by $C_i(t) = t + \theta_i, \ i = 1, \ldots, M$, where θ_i denotes the unknown clock offset of the ith anchor clock $C_i(t)$ relative to the absolute clock.

To make full use of the broadcast property of wireless signals, we employ the ATR protocol proposed in [136, 137] to make all the other anchors listen to
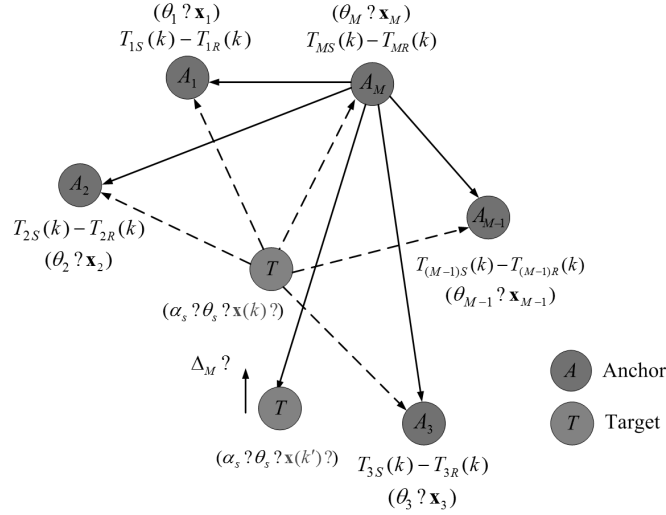
Figure 8.1: An example of the ATR protocol for mobile asynchronous networks

the ranging packets and record their timestamps locally, when one anchor and the target node exchange their ranging packets. This way, anchors obtain more information than for the two-way ranging (TWR) protocol proposed in the IEEE 802.15.4a standard [75] without increasing the communication load. The same packet structure as used in the standard is employed here, which is composed of a synchronization header (SHR) preamble, a physical layer header (PHR) and a data field. The first pulse of the PHR is called the ranging marker (RMARKER). The moment when the RMARKER leaves or arrives at the antenna of a node is critical to ranging. Without loss of generality, we assume that the Mth anchor initiates the ATR protocol as illustrated in Fig. 8.1. The ith anchor records the timestamps $T_{iR}(k)$ and $T_{iS}(k)$ upon the arrival of the RMARKERs of the ranging request from the Mth anchor and of the ranging response from the target node, respectively, where k is a label to indicate that the timestamp measurements correspond to $\mathbf{x}(k)$. Note that $T_{MR}(k)$ can be interpreted as the time upon which the Mth anchor receives its own ranging request without any delay, and it is recorded when the Mth anchor transmits a ranging packet. Because we do not use any timestamps from the target node, the clock parameters of the target node do not have any influence on our scheme. This is an important advantage of the ATR protocol compared to the TWR protocol. For the ith anchor, the difference between $T_{iR}(k)$ and $T_{iS}(k)$ relates to the distance as

$$c(T_{iS}(k) - T_{iR}(k)) = d_i(k) + d_M(k') + \Delta(k) - d_{iM} + n_{iS}(k) - n_{iR}(k),$$
$$i = 1, \ldots, M, \quad (8.1)$$

where $k' = k - \Delta(k)/c$, c is the signal propagation speed, $\Delta(k)$ is the unknown distance corresponding to the processing time of the target node, $d_i(k) = \|\mathbf{x}_i - \mathbf{x}(k)\| = \sqrt{\|\mathbf{x}_i\|^2 - 2\mathbf{x}_i^T\mathbf{x}(k) + \|\mathbf{x}(k)\|^2}$ is unknown, and $d_{ij} = \|\mathbf{x}_i - \mathbf{x}_j\|$ is known. Since the target node is moving continuously, the position where it receives the RMARKER from the Mth anchor is different from the position where it sends out its RMARKER, and the time interval is the processing time $\Delta(k)/c$. Thus, the distance between the target node and the Mth anchor is $d_M(k')$, when the target node receives the RMARKER from the Mth anchor. As a result, there are two unknown target positions $\mathbf{x}(k)$ and $\mathbf{x}(k')$ in (8.1). Furthermore, $n_{iS}(k)$ and $n_{iR}(k)$ are the distance error terms translated from the measurement errors in $T_{iS}(k)$ and $T_{iR}(k)$, which can be modeled as zero mean random variables with variance $\sigma_{iS}^2(k)$ and $\sigma_{iR}^2(k)$, respectively. By making differences of the timestamps from the same anchor, the clock offsets are canceled out. Moreover, the internal delays of all the anchors except the Mth anchor are also eliminated, since both $T_{iR}(k)$ and $T_{iS}(k)$ are recorded upon the arrival of the RMARKERS at the same node. The internal delay of the Mth anchor can be compensated beforehand [151]. Consequently, defining $\mathbf{q}(k) = c[T_{1S}(k) - T_{1R}(k), \ldots, T_{MS}(k) - T_{MT}(k)]^T$, $\mathbf{d}(k) = [d_1(k), \ldots, d_{M-1}(k), d_M(k)]^T$, $\mathbf{d}_a = [d_{1M}, \ldots, d_{(M-1)M}, 0]^T$, $\mathbf{n}_s = [n_{1S}(k), \ldots, n_{MS}(k)]^T$, and $\mathbf{n}_r = [n_{1R}(k), \ldots, n_{MR}(k)]^T$, we can now write (8.1) in vector form as

$$\mathbf{q}(k) = \mathbf{d}(k) + (d_M(k') + \Delta(k))\mathbf{1}_M - \mathbf{d}_a + \mathbf{n}_s(k) - \mathbf{n}_r(k). \qquad (8.2)$$

As (8.2) is a nonlinear model w.r.t. $\mathbf{x}(k)$ and $\mathbf{x}(k')$, it is impossible to derive the conventional KF for (8.2). Inspired by [166], we would like to linearize (8.2) without any approximation by projection and element-wise multiplication. We employ an orthogonal projection $\mathbf{P}$ onto the orthogonal complement of $\mathbf{1}_M$ similarly as in [137], which is defined as $\mathbf{P} = \mathbf{I}_M - \frac{1}{M}\mathbf{1}_M\mathbf{1}_M^T$. Since $\mathbf{P}\mathbf{1}_M = \mathbf{0}_M$, $\mathbf{P}$ can be used to eliminate the term $(d_M(k') + \Delta(k))\mathbf{1}_M$ in (8.2). As a result, premultiplying both sides of (8.2) with $\mathbf{P}$, we obtain

$$\mathbf{P}\mathbf{q}(k) = \mathbf{d}(k) - \bar{d}(k)\mathbf{1}_M - \mathbf{P}\mathbf{d}_a + \mathbf{P}\mathbf{n}_s(k) - \mathbf{P}\mathbf{n}_r(k), \qquad (8.3)$$

where $\mathbf{P}\mathbf{d}(k) = \mathbf{d}(k) - \bar{d}(k)\mathbf{1}_M$ with $\bar{d}(k) = \frac{1}{M}\sum_{i=1}^{M} d_i(k)$ being the unknown average of the distances between the target node and the anchors. Note that (8.3) is now only related to $\mathbf{x}(k)$ with the penalty of losing some information due to the projection. Keeping $\mathbf{d}(k)$ on one side, moving the other terms to the other side, and making an element-wise multiplication, we can write

$$\boldsymbol{\psi}_a - 2\mathbf{X}_a^T\mathbf{x}(k) + \|\mathbf{x}(k)\|^2\mathbf{1}_M = (\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)) \odot (\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a))$$
$$+ \bar{d}^2(k)\mathbf{1}_M + 2\bar{d}(k)\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a) + \mathbf{n}(k),$$
$$(8.4)$$

where $\boldsymbol{\psi}_a = [\|\mathbf{x}_1\|^2, \ldots, \|\mathbf{x}_M\|^2]^T$, and $\mathbf{n}(k) = -(\mathbf{P}(\mathbf{n}_s(k) - \mathbf{n}_r(k))) \odot (\mathbf{P}(\mathbf{n}_s(k) -$

$\mathbf{n}_r(k))) - 2\mathbf{d}(k) \odot \mathbf{P}(\mathbf{n}_s(k) - \mathbf{n}_r(k))$ with $\odot$ denoting element-wise product. Since the unconstrained least squares (LS) estimation method is equivalent to the subspace minimization (SM) method [123], we employ the latter one in order to estimate $\mathbf{x}(k)$ alone. We first apply $\mathbf{P}$ again to eliminate $\|\mathbf{x}(k)\|^2 \mathbf{1}_M$ and $\bar{d}^2(k)\mathbf{1}_M$, leading to

$$
\begin{aligned}
&\mathbf{P}\boldsymbol{\psi}_a - \mathbf{P}((\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)) \odot (\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a))) \\
=\ &2\mathbf{P}\mathbf{X}_a^T\mathbf{x}(k) + 2\bar{d}(k)\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a) + \mathbf{P}\mathbf{n}(k).
\end{aligned}
\tag{8.5}
$$

We then apply an orthogonal projection $\mathbf{P}_d(k)$ onto the orthogonal complement of $\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)$ to both sides of (8.5), which is given by

$$
\mathbf{P}_d(k) = \mathbf{I}_M - \frac{\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)(\mathbf{q}(k) + \mathbf{d}_a)^T\mathbf{P}}{(\mathbf{q}(k) + \mathbf{d}_a)^T\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)}.
\tag{8.6}
$$

As a result, we arrive at

$$
\mathbf{b}(k) = \mathbf{F}(k)\mathbf{x}(k) + \mathbf{P}_d(k)\mathbf{P}\mathbf{n}(k),
\tag{8.7}
$$

where $\mathbf{b}(k) = \mathbf{P}_d(k)\mathbf{P}\boldsymbol{\psi}_a - \mathbf{P}_d(k)\mathbf{P}((\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)) \odot (\mathbf{P}(\mathbf{q}(k) + \mathbf{d}_a)))$ and $\mathbf{F}(k) = 2\mathbf{P}_d(k)\mathbf{P}\mathbf{X}_a^T$. Note that $\mathbf{P}_d(k)$, $\mathbf{b}(k)$ and $\mathbf{F}(k)$ all depend on time-varying measurements. We remark that in order to facilitate all the linearizations, the condition $M \geq l + 3$ has to be fulfilled, which indicates that we need at least five anchors on a plane or six anchors in space to accomplish the linearization.

Let us now explore the statistical properties of the noise. Defining $\mathbf{P}\mathbf{n}_r(k) = \mathbf{n}_r(k) - \bar{n}_r(k)\mathbf{1}_M$ and $\mathbf{P}\mathbf{n}_s(k) = \mathbf{n}_s(k) - \bar{n}_s(k)\mathbf{1}_M$, where $\bar{n}_r(k) = \frac{1}{M}\sum_{i=1}^M n_{iR}(k)$ and $\bar{n}_s(k) = \frac{1}{M}\sum_{i=1}^M n_{iS}(k)$, we can write the entries of $\mathbf{n}(k)$ as

$$
\begin{aligned}
[\mathbf{n}(k)]_i =\ &2d_i(n_{iR}(k) - \bar{n}_r(k) - n_{iS}(k) + \bar{n}_s(k)) \\
&-(n_{iR}(k) - \bar{n}_r(k) - n_{iS}(k) + \bar{n}_s(k))^2, i = 1, 2, \ldots, M.
\end{aligned}
\tag{8.8}
$$

Recall that $E[n_{iS}(k)] = 0$, $E[n_{iS}^2(k)] = \sigma_{iS}^2(k)$ and $E[n_{iS}(k)n_{jS}(k)] = 0, i \neq j$, which leads to $E[\bar{n}_s(k)] = 0$, $E[\bar{n}_s^2(k)] = \frac{1}{M^2}\sum_{i=1}^M \sigma_{iS}^2(k)$ and $E[\bar{n}_s(k)n_{iS}(k)] = \frac{1}{M}\sigma_{iS}^2(k)$. The statistical properties of $n_{iR}(k)$ can be obtained in a similar way. Moreover, $n_{iS}(k)$ and $n_{iR}(k), i = 1, \ldots, M$ are uncorrelated. As a result, the statistical

properties of $\mathbf{n}(k)$ are given by

$$E[[\mathbf{n}(k)]_i]$$
$$= \frac{2-M}{M}(\sigma_{iR}^2(k) + \sigma_{iS}^2(k)) - \frac{1}{M^2}\sum_{p=1}^{M}(\sigma_{pR}^2(k) + \sigma_{pS}^2(k)) \approx 0, \qquad (8.9)$$

$$E[[\mathbf{n}(k)]_i[\mathbf{n}(k)]_j]$$
$$\approx \begin{cases} 4d_i(k)^2 \left(\frac{M-2}{M}(\sigma_{iR}^2(k) + \sigma_{iS}^2(k)) + \frac{1}{M^2}\sum_{p=1}^{M}(\sigma_{pR}^2(k) + \sigma_{pS}^2(k)) \right), i = j, \\ 4d_i(k)d_j(k) \left(\frac{1}{M^2}\sum_{p=1}^{M}(\sigma_{pS}^2(k) + \sigma_{pR}^2(k)) \right. \\ \left. -\frac{1}{M}(\sigma_{iS}^2(k) + \sigma_{jS}^2(k) + \sigma_{iR}^2(k) + \sigma_{jR}^2(k)) \right), i \neq j, \end{cases}$$
$$(8.10)$$

where we ignore the higher order noise terms to obtain (8.10) and assume $E[[\mathbf{n}(k)]_i] \approx 0$ under the condition of sufficiently small measurement errors. Thus, we still treat $\mathbf{n}(k)$ as a zero mean Gaussian random vector. We remark that the noise covariance matrix depends on the unknown $\mathbf{d}(k)$. To solve this problem, we can plug in the predicted $\hat{\mathbf{d}}(k|k-1)$, which makes use of the prediction $\hat{\mathbf{x}}(k|k-1)$ (these notations will be defined later on). Note that $\mathbf{n}(k)$ is not a stationary process but it is independent, leading to

$$\mathbf{P}_d(k)\mathbf{P}E[\mathbf{n}(k)\mathbf{n}(j)^T]\mathbf{P}\mathbf{P}_d(j) = \begin{cases} \mathbf{\Lambda}(k) & k = j \\ \mathbf{0} & k \neq j \end{cases}, \qquad (8.11)$$

where $\mathbf{\Lambda}(k)$ is rank-deficient due to the projections.

8.3   Dynamic state model and Kalman filter

Let us define the state at time k as $\mathbf{s}(k) = [\mathbf{x}(k)^T, \ \dot{\mathbf{x}}(k)^T, \ \ddot{\mathbf{x}}(k)^T]^T$, where $\mathbf{x}(k)$, $\dot{\mathbf{x}}(k)$ and $\ddot{\mathbf{x}}(k)$ are the coordinate, the velocity and the acceleration vectors of the target node at time k, respectively. We assume a general linear state model, which is given by (see also [25])

$$\mathbf{s}(k+1) = \mathbf{A}(k)\mathbf{s}(k) + \mathbf{B}(k)\mathbf{u}(k) + \mathbf{w}(k), \qquad (8.12)$$

where $\mathbf{A}(k)$ is a $3l \times 3l$ state transition matrix, $\mathbf{B}(k)$ is a $3l \times l$ input matrix, $\mathbf{u}(k)$ is an acceleration input vector of length l, and $\mathbf{w}(k)$ is a driving noise vector of length l with zero mean and a covariance matrix $\mathbf{R}(k)$, which is given by

$$E[\mathbf{w}(k)\mathbf{w}(j)^T] = \begin{cases} \mathbf{R}(k) & k = j \\ \mathbf{0} & k \neq j \end{cases}. \qquad (8.13)$$

We assume that $\mathbf{A}(k)$, $\mathbf{B}(k)$ and $\mathbf{u}(k)$ are all known exactly. Moreover, the driving noise and the measurement noise are assumed independent. In practice,

$\mathbf{u}(k)$ has to be estimated first. In [167] for instance, $\mathbf{u}(k)$ is modeled as a semi-Markov process with a finite number of possible acceleration inputs, which are selected based on the transition probabilities of a Markov process. As a result, $\mathbf{u}(k)$ can be estimated by a minimum mean square error (MMSE) estimator. On the other hand, [168] does not require a statistical model for $\mathbf{u}(k)$, and derives a LS estimator of $\mathbf{u}(k)$.

Let us now rewrite the measurement model (8.7) using $\mathbf{s}(k)$ as

$$\mathbf{b}(k) = \mathbf{C}(k)\mathbf{s}(k) + \mathbf{P}_d(k)\mathbf{Pn}(k), \tag{8.14}$$

where $\mathbf{C}(k) = [\mathbf{F}(k),\ \mathbf{0}_{M \times 2l}]$. Based on (8.12) and (8.14), we can easily develop the corresponding KF tracker. The prediction equations are given by

$$\hat{\mathbf{s}}(k|k{-}1) = \mathbf{A}(k{-}1)\hat{\mathbf{s}}(k{-}1|k{-}1) + \mathbf{B}(k{-}1)\mathbf{u}(k{-}1), \tag{8.15}$$

$$\mathbf{P}_s(k|k{-}1) = \mathbf{A}(k{-}1)\mathbf{P}_s(k{-}1|k{-}1)\mathbf{A}(k{-}1)^T + \mathbf{R}(k{-}1). \tag{8.16}$$

The update equations are given by

$$\mathbf{K}(k) = \mathbf{P}_s(k|k-1)\mathbf{C}(k)^T \left(\mathbf{C}(k)\mathbf{P}_s(k|k-1)\mathbf{C}(k)^T + \mathbf{\Lambda}(k)\right)^\dagger, \tag{8.17}$$

$$\hat{\mathbf{s}}(k|k) = \hat{\mathbf{s}}(k|k-1) + \mathbf{K}(k)\left(\mathbf{b}(k) - \mathbf{C}(k)\hat{\mathbf{s}}(k|k-1)\right), \tag{8.18}$$

$$\mathbf{P}_s(k|k) = \left(\mathbf{I}_M - \mathbf{K}(k)\mathbf{C}(k)\right)\mathbf{P}_s(k|k-1). \tag{8.19}$$

Note that since $\mathbf{C}(k)\mathbf{P}_s(k|k-1)\mathbf{C}(k)^T + \mathbf{\Lambda}(k)$ may be rank-deficient, we use the pseudo-inverse instead of the inverse, which is denoted here by $(\cdot)^\dagger$.

8.4   Extended Kalman filter

In this section, we derive the EKF as a benchmark for our KF tracker. In order to apply the EKF, we first have to use a Taylor expansion to linearize the non-linear measurement model. Recall the data model (8.3) here, which is the result of applying the projection $\mathbf{P}$ to (8.2) in order to get rid of the dominant term $(d_M(k') + \Delta(k))\mathbf{1}_M$, where $\Delta(k)$ is the unknown distance corresponding to the processing time, and $d_M(k')$ is the distance between $\mathbf{x}(k')$ and the Mth anchor. As a result of the projection, (8.3) is only related to $\mathbf{x}(k)$:

$$\mathbf{Pq}(k) = \mathbf{d}(k) - \bar{d}(k)\mathbf{1}_M - \mathbf{Pd}_a + \mathbf{P}(\mathbf{n}_s(k) - \mathbf{n}_r(k)), \tag{8.20}$$

where we recall that $\bar{d}(k) = \frac{1}{M}\sum_{i=1}^{M} d_i(k)$. Let us define the function $f(\mathbf{x}(k))$ as $f(\mathbf{x}(k)) = \mathbf{d}(k) - \bar{d}(k)\mathbf{1}_M - \mathbf{Pd}_a(k)$, where we recall that $d_i(k) = \|\mathbf{x}(k) - \mathbf{x}_i(k)\|$.

The Jacobian $\mathbf{H}(k)$ of $f(\mathbf{x}(k))$ w.r.t. $\mathbf{s}(k)$ can be expressed as

$$[\mathbf{H}(k)]_{i,j} = \left.\frac{\partial[f(\mathbf{x}(k))]_i}{\partial[\mathbf{s}(k)]_j}\right|_{\mathbf{s}(k)=\hat{\mathbf{s}}(k|k-1)}, \quad \text{with}$$

$$\left.\frac{\partial[f(\mathbf{x}(k))]_i}{\partial\mathbf{x}(k)}\right|_{\mathbf{x}(k)=\hat{\mathbf{x}}(k|k-1)} = \left(\frac{(\mathbf{x}(k)-\mathbf{x}_i(k))^T}{\|\mathbf{x}(k)-\mathbf{x}_i(k)\|}\right. \tag{8.21}$$

$$\left.-\frac{1}{M}\sum_{j=1}^{M}\frac{(\mathbf{x}(k)-\mathbf{x}_j(k))^T}{\|\mathbf{x}(k)-\mathbf{x}_j(k)\|}\right)\Bigg|_{\mathbf{x}(k)=\hat{\mathbf{x}}(k|k-1)} \tag{8.22}$$

$$\tag{8.23}$$

$$\left.\frac{\partial[f(\mathbf{x}(k))]_i}{\partial\dot{\mathbf{x}}(k)}\right|_{\dot{\mathbf{x}}(k)=\hat{\dot{\mathbf{x}}}(k|k-1)} = \left.\frac{\partial[f(\mathbf{x}(k))]_i}{\partial\ddot{\mathbf{x}}(k)}\right|_{\ddot{\mathbf{x}}(k)=\hat{\ddot{\mathbf{x}}}(k|k-1)} = \mathbf{0}_l^T. \tag{8.24}$$

Recall that $\mathbf{n}_s(k)$ and $\mathbf{n}_r(k)$ are zero mean independent Gaussian random variables with variance $\sigma_{iS}^2(k)$ and $\sigma_{iR}^2(k)$, respectively. Defining $\widetilde{\mathbf{\Lambda}}(k)$ as the noise covariance matrix of the noise term $\mathbf{P}(\mathbf{n}_s(k)-\mathbf{n}_r(k))$, which is a zero mean Gaussian random vector, we can write $\widetilde{\mathbf{\Lambda}}(k)=\mathbf{P}\text{diag}([\sigma_{1S}^2(k)+\sigma_{1R}^2(k), \ldots, \sigma_{MS}^2(k)+\sigma_{MR}^2(k)])\mathbf{P}$.

Consequently, the EKF is developed as follows. The prediction equations are the same as (8.15) and (8.16). The update equations are

$$\widetilde{\mathbf{K}}(k)=\widetilde{\mathbf{P}}_s(k|k-1)\mathbf{H}(k)^T(\mathbf{H}(k)\widetilde{\mathbf{P}}_s(k|k-1)\mathbf{H}(k)^T+\widetilde{\mathbf{\Lambda}}(k))^\dagger, \tag{8.25}$$

$$\hat{\tilde{\mathbf{s}}}(k|k)=\hat{\tilde{\mathbf{s}}}(k|k-1)+\widetilde{\mathbf{K}}(k)(\mathbf{P}\mathbf{q}(k)-f(\hat{\mathbf{x}}(k|k-1))), \tag{8.26}$$

$$\widetilde{\mathbf{P}}_s(k|k)=(\mathbf{I}_M-\widetilde{\mathbf{K}}(k)\mathbf{H}(k))\widetilde{\mathbf{P}}_s(k|k-1). \tag{8.27}$$

8.5   Simulation results

Let us now evaluate the performance of the proposed KF tracker by Monte Carlo simulations, and compare it with the EKF. We consider a simulation setup, where the first anchor is located at the origin, and the other four anchors are located at the corners of a $100\,\text{m}\times100\,\text{m}$ rectangular centered around the origin. Due to the broadcast nature of the ATR protocol, we assume that $\sigma_{iS}^2(k)$ and $\sigma_{iR}^2(k)$ are related to the distances according to the path loss law. Thus we define the average noise power as $\bar{\sigma}^2 = 1/M\sum_{i=1}^{M}\sigma_{iS}^2(k)$, where $\sigma_{iR}^2(k)$ and $\sigma_{iS}^2(k)$ are chosen to fulfill the condition that all $\sigma_{iR}^2(k)/d_{iM}^2$ and $\sigma_{iS}^2(k)/d_i^2(k)$ are equal. Note that since $d_{MM}=0$, we simply assume $\sigma_{MR}^2(k)=0$ and $n_{MR}=0$. The processing time of the target node is 5 ms, and the signal propagation speed c is the speed of the light. As a result the corresponding distance $\Delta(k)$ is $3\times10^8\times5\times10^{-3}=1.5\times10^6$ m, which is much larger than the scale of the considered set-up. We employ a
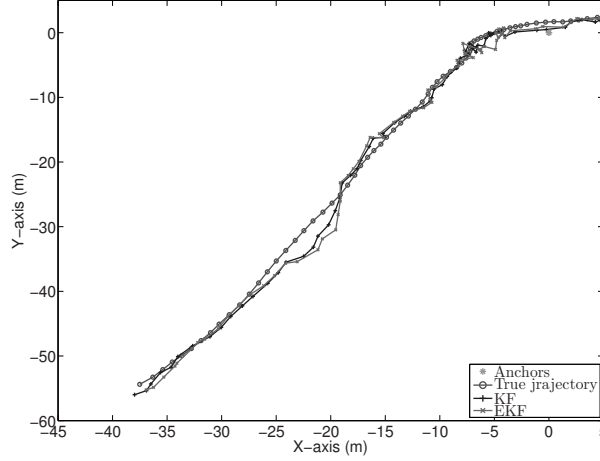
Figure 8.2: An example of the trajectories estimated by the proposed KF and the EKF

random walk state model as in [57], where (8.12) is reduced to

$$\mathbf{s}(k+1) = \left[\begin{array}{ccc} \mathbf{I}_l & T_s\mathbf{I}_l & \mathbf{0}_{l\times l} \\ \mathbf{0}_{l\times l} & \mathbf{I}_l & \mathbf{0}_{l\times l} \\ \mathbf{0}_{l\times l} & \mathbf{0}_{l\times l} & \mathbf{0}_{l\times l} \end{array}\right]\mathbf{s}(k) + \left[\begin{array}{c} \mathbf{0}_l \\ \mathbf{w}_{\dot{\mathbf{x}}}(k) \\ \mathbf{0}_l \end{array}\right], \tag{8.28}$$

where $\mathbf{w}_{\dot{\mathbf{x}}}(k)$ is a zero mean white random process with covariance matrix $\sigma_w^2\mathbf{I}_l$, and T_s is the sampling interval. For a true initial state $\mathbf{s}(-1) = [\mathbf{x}(-1)^T, \ \mathbf{v}(-1)^T, \ \mathbf{0}_l^T]^T$, the initial state estimate $\hat{\mathbf{s}}(-1|-1)$ is randomly generated in each Monte Carlo run according to $\mathcal{N}(\mathbf{s}(-1), \mathbf{P}_s(-1|-1))$, with $\mathbf{P}_s(-1|-1) = 100\text{diag}([\mathbf{1}_l^T, \ \mathbf{1}_l^T, \ \mathbf{0}_l^T]^T)$. The same initial state estimate is also used for the EKF. In each run, we generate a trajectory of 100 points based on the state model. The performance criterion is the root mean square error (RMSE) of $\hat{\mathbf{x}}$ (or $\hat{\dot{\mathbf{x}}}$) vs. the time index k, which can be expressed as $\sqrt{1/N_{exp}\sum_{j=1}^{N_{exp}} \|\hat{\mathbf{x}}^{(j)} - \mathbf{x}\|^2}$, where $\hat{\mathbf{x}}^{(j)}$ is the estimate obtained in the jth trial. Each simulation result is averaged over $N_{exp} = 500$ Monte Carlo trials. The rest of the parameters are given by $\mathbf{x}(-1) = [13\,\text{m}, \ 4\,\text{m}]^T$, $\mathbf{v}(-1) = [0.05\,\text{m/s}, \ -0.05\,\text{m/s}]^T$, $T_s = 1\,\text{s}$, $\sigma_{\mathbf{x}}^2 = \sigma_{\dot{\mathbf{x}}}^2 = \sigma_w^2 = 0.01$, $1/\bar{\sigma}^2 = 20\,\text{dBm}$.

Fig. 8.2 shows an example of the trajectories estimated by the proposed KF (the line with "+" markers) and the EKF (the line with "×" markers), respectively. The true trajectory is the line with "∘" markers. Fig. 8.3(a) illustrates the performance of the position estimate $\hat{\mathbf{x}}$ by the proposed KF and the EKF, respectively. The proposed KF achieves better accuracy than the EKF. Both
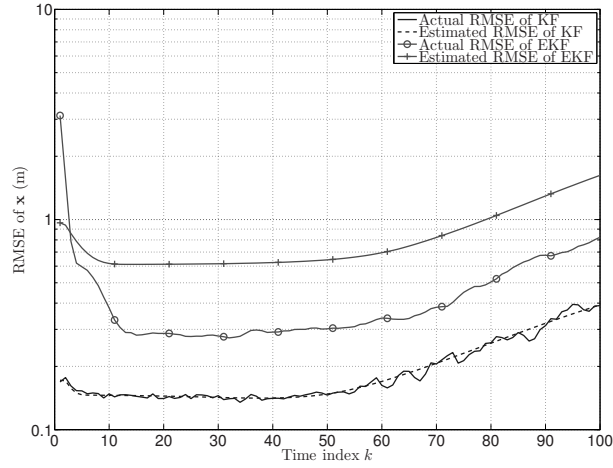
(a) RMSE of the target node location $\mathbf{x}$ vs. time index k
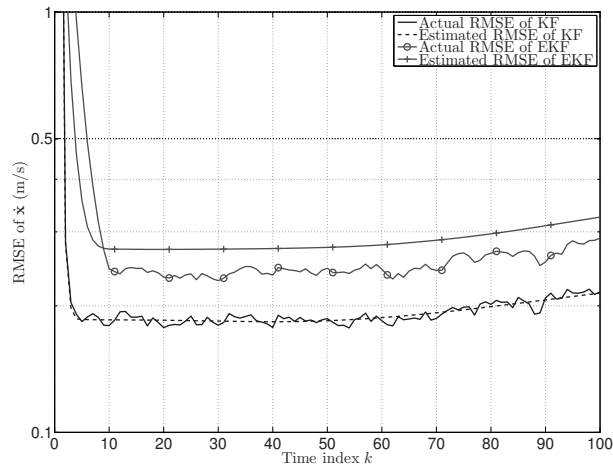
(b) RMSE of the target node velocity $\dot{\mathbf{x}}$ vs. time index k

Figure 8.3: RMSE of the proposed KF and the EKF

RMSEs increase with time, since the covariance matrix of the state also increases with time. Furthermore, the true RMSE (the solid line) of the proposed KF closely follows its estimated RMSE (the dashed line), which is calculated by $\sqrt{1/N_{exp} \sum_{j=1}^{N_{exp}} ([\mathbf{P}^{(j)}(k|k)]_{1,1} + [\mathbf{P}^{(j)}(k|k)]_{2,2})}$ with $\mathbf{P}^{(j)}(k|k)$ the covariance estimate obtained in the jth trial. On the other hand, the true RMSE (the solid line

with "∘" markers) of the EKF is below its estimated RMSE (the solid line with "+" markers), which shows that the estimated RMSE of the EKF is pessimistic. We observe the same tendency in Fig. 8.3(b), which illustrates the performance of the velocity estimate $\hat{\dot{\mathbf{x}}}$ by the proposed KF and the EKF, respectively.

8.6   Conclusions

In this chapter, a KF based tracking method is developed to track the target node with the help of anchors in asynchronous networks with clock offsets. Since we make difference of the TOA measurements obtained by the same anchor, the clock offsets are canceled out, and the influence of the asynchronous target clock is avoided. Moreover, a linear measurement model is derived. Thus, the proposed KF based on this exact linear measurement model does not have the modeling errors inherently contained in the EKF. Furthermore, low computational complexity makes the proposed KF a promising solution for practical use. According to the simulation results, the proposed KF outperforms the EKF. In future work, we would like to propose low complexity trackers for asynchronous WSNs not only with clock offsets but also with clock skews.

CHAPTER **9**

Conclusions and Future Work

9.1 Conclusions

The research presented in this thesis is simulated by the great demand for location-aware wireless sensor networks (WSNs). The unique characteristics of WSNs impose numerous challenges on localization and communication. In this thesis, we tackle some key challenges and provide affordable solutions. Impulse radio ultra wideband (IR-UWB) is employed as the backbone technology for both localization and communication due to its distinctive advantages in ranging and reliable communication. Moreover, synchronization is the red thread running through the whole thesis. We review the main results of the thesis in this section, and discuss possible future work in the next section.

We started with a transmitted-reference (TR) UWB communication system. The TR-UWB scheme obtains sub-Nyquist rate samples by correlating the received pulse sequence with its delayed version in the analog domain. Thus, it avoids the daunting Nyquist sampling rate, relaxes the stringent synchronization requirements, and only asks for aggregate channel coefficients. We aimed at a moderate data rate for WSNs, then adopted a data model taking various kinds of interferences into account, and proposed a complete solution for signal detection, channel estimation, synchronization and equalization. We found that the simple detector, which sums up all the samples in the observation window and compares the result with a threshold, gives a good balance between the detection performance and the implementation cost. Moreover, a two-stage synchronization strategy was proposed to first achieve sample-level synchronization and later symbol-level synchronization. Different combinations of channel estim-

ation and equalization schemes were investigated using the IEEE UWB channel model (IEEE 802.15.3a, CM3, none-line-of-sight (NLOS, 4-10 m) channels). In our simulation setup (with a data rate of 2.2 Mb/s), the optimal combination considering cost and performance would be a matched filter (MF) channel estimator with a zero-forcing (ZF) equalizer. All the above results indicate that the TR-UWB system can be implemented with low cost and obtain moderate data rate communications.

Next, we investigated the theoretical ranging accuracy of a novel method, which exploits the range information in both the amplitude and the time delay of the received signal. Several bounds were derived not only in an additive white Gaussian noise (AWGN) channel with attenuation, but also in an AWGN channel with both attenuation and shadowing. For UWB signals, the proposed novel method does not show obvious benefits. However, it indeed improves the ranging accuracy using narrow band (NB) signals. Moreover, taking the prior knowledge of the shadowing effects into account can further lower the bounds for range estimation.

As we have already explored the theoretical bounds for ranging, we continued to develop a practical ranging method. We employed the stroboscopic sampling for an IR-UWB system to achieve accurate time-of-arrival (TOA) estimation for ranging. Hence, we avoided Nyquist sampling at the cost of transmission efficiency, but maintained the same ranging resolution as Nyquist sampling can achieve. Due to the long preamble required by stroboscopic sampling, the clock drift, which is an accumulative effect over time caused by the relative clock skew between different clocks, is one of the main error sources in TOA estimation. Therefore, we proposed TOA estimation methods with clock drift calibration, which dramatically reduces the TOA estimation errors due to the drift. As a result, we have proposed a practical low sampling rate solution for accurate TOA estimation using IR-UWB.

Since we developed all ingredients for ranging fusion, we further investigated low-complexity methods to fuse all range measurements for location estimation in WSNs. We first considered a scenario, where no anchors (sensors with known positions) are available, and our goal is to jointly estimate the positions of all the nodes given partial pairwise distance measurements up to a translation, rotation, and reflection. We extended the classical multi-dimensional scaling (MDS) algorithm by two projections to handle networks composed of two groups: the first group contains nodes that communicate with each other, and the second group contains nodes that do not communicate with each other, but only communicate with each of the nodes in the first group. As there are missing links in this kind of networks, the classical MDS cannot be applied. Our approach outperforms other methods based on matrix completion techniques.

Note that in Chapter 5 we have assumed that the distance measurements are available, and do not care about how to obtain them. The ignorance imposes great challenges on the ranging step. Therefore, inspired by the cross-layer philosophy, we relax the strict separation between the ranging and the ranging fusion,

explicitly model ranging errors and handle them in the ranging fusion step. It leads us to a new way to deal with the localization problem. In the rest part of the thesis, we considered a scenario, where a target sensor node is connected with many anchors. We proposed reference-free localization estimators based on TOA measurements for networks, where anchors are synchronized, and the clock of the target node runs freely. The ranging bias due to the asynchronous target clock was explicitly taken into account in the data model. For such kind of networks, a reference anchor is conventionally selected to get rid of this bias, or time-difference-of-arrival (TDOA) measurements are directly obtained, where this bias is canceled out inherently. However, a poor reference selection can seriously degrade the localization performance of reference-based least-squares (LS) estimators. On the other hand, the reference-free estimators do not suffer from it. In this thesis, we also generalized existing reference-based closed-form localization estimators using TOA or TDOA measurements, and exposed their relations. We obtained several important conclusions based on analysis and simulations: (i) applying a projection is always preferred over making differences with a reference to get rid of nuisance parameters; (ii) the optimal weighting matrix can compensate for the influence of the reference selection for reference-based WLS estimators using TOA measurements, however it cannot decouple the reference dependency for WLS estimators using a nonredundant set of TDOA measurements; (iii) there are corresponding equivalences between the squared-range-based and the squared-range-difference-based methods, which are all using TOA measurements; (iv) beyond some SNR threshold, there are no obvious differences among the CRBs using TOA measurements, the nonredundant set and the full set of TDOA measurements, respectively; (v) the performance of the reference-free LS estimators is neither too bad nor too good, but they do not suffer from a poor reference selection; and (vi) the concrete value of the distance bias caused by the inaccurate clock does not affect the localization performance of the LS or WLS estimators.

Sequentially, we proposed robust localization strategies based on TOA measurements to localize a target node with the help of anchors for asynchronous networks. We dealt with two kinds of asynchronous networks with respect to (w.r.t) anchor clocks: one only with clock offsets referred to as quasi-synchronous networks, and the other not only with clock offsets but also with clock skews referred to as fully asynchronous networks. Note that the target clock can always run freely in both networks. We take the clock parameters into our data model to jointly estimate clock parameters and target position. Furthermore, regardless of the reliability of the timestamps from the target node, we proposed a novel ranging protocol, namely asymmetric trip ranging (ATR), which leads to localization methods that are naturally immune to internal attacks mounted by a compromised target node. Compared to the conventional two-way ranging (TWR) protocol, the ATR protocol makes use of the broadcast property of wireless networks, and has a smaller communication load. Closed-form LS and iterative WLS estimators were proposed to localize the target node for quasi-synchronous networks. For fully

asynchronous networks, we developed two different approaches with low complexity to jointly estimate the position and clock parameters. The one with the minimum communication load takes practical issues into account to simplify the estimators by appropriate approximation. The other relaxes the communication load constraint in order to obtain accurate estimates. We can choose between them according to system specifications. Furthermore, we designed closed-form LS estimators for joint synchronization and localization using the TWR protocol, and compared them with the proposed closed-form estimators employing the ATR protocol.

Last but not least, a Kalman filter (KF) based tracking method was developed to track the target node with the help of anchors in quasi-synchronous networks defined in the previous paragraph. The ATR protocol was employed again to facilitate clock offset cancelation. A linear measurement model based on TOA measurements was derived by projection and element-wise multiplication. Hence, the proposed KF based on this exact linear measurement model does not have the modeling errors inherently contained in the extended KF (EKF). Moreover, the low computational complexity makes the proposed KF a promising solution for practical use.

The Smart PEAS project is a good application for the above results. We have developed a prototype system (see Fig. 1.5 in Chapter 1) together with 3UB (a TU Delft spin-off company) and Delft Engineering Services (DES) for the Smart PEAS project, which aims at investigating the use of smart moving process environment actuators and sensors (PEAS) based on UWB technology inside the process equipment for product quality control. An aquarium (1 m × 1 m × 1 m) full of water is used in the prototype system to mimic a process environment. A mechanical system precisely controls the target sensor position. The ball-shaped PEAS (see Fig. 1.6 in Chapter 1) are designed with a bow-tie antenna on the surface and the sensor electronics embedded inside. Several anchors are employed and directly connected to high speed scopes to capture the signals. Online localization using this platform is currently under test.

9.2   Future work

The adventure of pursuing a PhD trains our skills, strengthens our capabilities, and matures our thoughts. Then, the completion of a PhD study opens the door to a more fascinating research world. In this section, we discuss a few possible directions for future work.

- Distributed localization

  The conventional localization methods for WSNs are centralized, where all the anchors and target sensors transmit their local measurements to a fusion center (FC). Location estimates for WSNs are computed at the FC. However, the centralized scheme has inherent limits: i) to transmit local in-

formation to a FC requires lots of resources, such as energy and bandwidth; and ii) the dependency on a single FC makes it the most powerful, but also the most vulnerable point in the whole network. Hence, the centralized scheme might not be well suitable for large scale WSNs. On the other hand, a distributed scheme gets rid of these limits, and provides a more robust solution. A distributed localization scheme for WSNs aims at obtaining globally optimal location estimates based on local information exchange. It iteratively refines local estimates maintained by each sensor node. The localization problem is difficult enough due to the nonlinear and nonconvex properties. Thus, it is even more challenging to solve it in a distributed fashion. To make full use of the convex optimization toolbox, the nonconvex localization problem was relaxed to a convex problem in [169, 170]. Moreover, quite a few distributed schemes were proposed for linear systems, such as consensus-based distribution methods [171–173], and distributed approaches via dual decomposition techniques [174]. Some research work formulates the localization problem in a distributed way, and optimizes the nonlinear and nonconvex cost function in each iteration. For example, the majorization method was employed in each iteration in [107], and second-order cone programming (SOCP) relaxation was applied in [175]. However, the computational complexity for each iteration is still quite high. As we have explored methods to linearize the localization problem in this thesis, there could be possibilities to obtain low-complexity linear localization estimators in a distributed fashion.

- Robust localization

  As we have mentioned in Chapter 1, the localization for WSNs can be accomplished sequentially. The sensors that reach enough anchors are localized first. Then, these located sensors can be viewed as new anchors that can facilitate the localization of other sensors. But the position estimates of these new anchors include estimation errors, which will propagate and eventually corrupt other estimates, if they are not taken into account. Furthermore, due to various kinds of error sources (e.g., NLOS, asynchronous clocks, uncertainty of the signal propagation speed or sensor failure), the range measurements may be contaminated by outliers. Even one outlier can break down all the estimates. Therefore, robust localization methods are in need to deal with these errors to obtain accurate estimates. Some pioneer work considers robust linear regression against outliers. The M-estimator exploiting the Huber cost function was proposed in [176]. Least median squares (LMS) was used as the criterion to counter outliers in [118, 177]. Furthermore, in [178, 179], outliers were explicitly modeled and an ℓ_1 norm regularization term was added to the standard LS criterion. Since the original localization problem is nonlinear, existing robust linear regression cannot be directly applied. Hence, we have to tailer our localization problem to take benefits from it.

- Tracking

  Sensor nodes in WSNs are often not static. They could be moving around. Tracking mobile sensors is definitely not trivial for the successful deployment of WSNs. Due to the nonlinear relation between the ranging (or bearing) measurements and the sensor positions, the conventional Kalman filter (KF) cannot be directly applied. Restricted by the cost and the power constraints of WSNs, computationally intensive methods, which deal with nonlinear models, such as the particle filter and the unscented KF (UKF), are not appropriate for WSNs. By linearizing the measurement model, we have revived the conventional KF. However, as long as the KF-like methods are employed, the measurement model and the state model have to be available. In practice, it is difficult to have an accurate state model, since there are always model mismatches [180–182]. Furthermore, inspired by the same reasons as for distributed localization, we would also like to implement tracking in a distributed fashion [183, 184]. Hence, robust distributed tracking approaches are of great interest for future work.

Bibliography

[1] S. Gezici, Z. Tian, G. Giannakis, H. Kobayashi, A. Molisch, H. Poor, and Z. Sahinoglu, "Localization via ultra-wideband radios: a look at positioning aspects for future sensor networks," *IEEE Signal Processing Mag.*, vol. 22, pp. 70–84, Jul. 2005.

[2] N. Patwari, J. Ash, S. Kyperountas, I. Hero, A.O., R. Moses, and N. Correal, "Locating the nodes: cooperative localization in wireless sensor networks," *IEEE Signal Processing Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.

[3] B. Parkinson and J. Spilker Jr., Eds., *Global Positioning System: Theory and Application, Volume I.* American Institute of Astronautics and Aeronautics, 1996.

[4] D. Torrieri, "Statistical theory of passive location systems," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 20, no. 2, pp. 183 –198, Mar. 1984.

[5] V. Chandrasekhar, W. K. Seah, Y. S. Choo, and H. V. Ee, "Localization in underwater sensor networks: survey and challenges," in *Proc. ACM WUWNet*, New York, USA, 2006, pp. 33–40.

[6] Y. Huang, J. Benesty, G. Elko, and R. Mersereati, "Real-time passive source localization: a practical linear-correction least-squares approach," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 9, no. 8, pp. 943–956, Nov. 2001.

[7] D. Li and Y. H. Hu, "Least square solutions of energy based acoustic source localization problems," in *Proc. ICPPW*, Montreal, Canada, Aug. 2004, pp. 443–446.

[8] J. Caffery, J. and G. Stuber, "Subscriber location in CDMA cellular networks," *IEEE Trans. Veh. Technol.*, vol. 47, no. 2, pp. 406 –416, May 1998.

[9] A. Sayed, A. Tarighat, and N. Khajehnouri, "Network-based wireless location: challenges faced in developing techniques for accurate wireless location information," *IEEE Signal Processing Mag.*, vol. 22, no. 4, pp. 24–40, Jul. 2005.

[10] H. Liu, H. Darabi, P. Banerjee, and J. Liu, "Survey of wireless indoor positioning techniques and systems," *IEEE Trans. Syst., Man, Cybern. C*, vol. 37, no. 6, pp. 1067 –1080, Nov. 2007.

[11] K. Cheung, H. So, W.-K. Ma, and Y. Chan, "Least squares algorithms for time-of-arrival-based mobile location," *IEEE Trans. Signal Processing*, vol. 52, no. 4, pp. 1121 – 1130, Apr. 2004.

[12] A. Savvides, C.-C. Han, and M. B. Strivastava, "Dynamic fine-grained localization in ad-hoc networks of sensors," in *Proc. ACM MobiCom*, Rome, Italy, Jul. 2001, pp. 166–179.

[13] D. Niculescu and B. Nath, "Ad hoc positioning system (APS) using AOA," in *Proc. IEEE INFOCOM*, vol. 3, San Francisco, CA, USA, Apr. 2003, pp. 1734 – 1743.

[14] P. Bahl and V. Padmanabhan, "RADAR: an in-building RF-based user location and tracking system," in *Proc. IEEE INFOCOM*, vol. 2, Tel Aviv, Israel, Mar. 2000, pp. 775 –784.

[15] R. Walden, "Analog-to-digital converter survey and analysis," *IEEE J. Select. Areas Commun.*, vol. 17, no. 4, pp. 539–550, Apr. 1999.

[16] B. Le, T. Rondeau, J. Reed, and C. Bostian, "Analog-to-digital converters," *IEEE Signal Processing Mag.*, vol. 22, no. 6, pp. 69–77, Nov. 2005.

[17] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *IEEE Trans. Signal Processing*, vol. 50, no. 6, pp. 1417 –1428, Jun. 2002.

[18] J. Paredes, G. Arce, and Z. Wang, "Ultra-wideband compressed sensing: Channel estimation," *IEEE J. Select. Topics Signal Processing*, vol. 1, no. 3, pp. 383 –395, Oct. 2007.

[19] G. Shi, J. Lin, X. Chen, F. Qi, D. Liu, and L. Zhang, "UWB echo signal detection with ultra-low rate sampling based on compressed sensing," *IEEE Trans. Circuits Syst. II*, vol. 55, no. 4, pp. 379 –383, Apr. 2008.

[20] D. Donoho, "Compressed sensing," *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289 –1306, Apr. 2006.

[21] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489 –509, Feb. 2006.

[22] A. Perrig, J. Stankovic, and D. Wagner, "Security in wireless sensor networks," *Commun. of the ACM*, vol. 47, no. 6, pp. 53–57, Jun. 2004.

[23] J. Walters, Z. Liang, W. Shi, and V. Chaudhary, "Wireless sensor network security: A survey," *Security in Distributed, Grid, and Pervasive Computing*, pp. 367 – 405, 2007.

[24] M. Flury, M. Poturalski, P. Papadimitratos, J.-P. Hubaux, and J.-Y. Le Boudec, "Effectiveness of distance-decreasing attacks against impulse radio ranging," in *Proc. ACM WiSec*, Hoboken, NJ, USA, Mar. 2010.

[25] Y. Shalom, X. Li, and K. T., *Estimation with applications to tracking and navigation.* New York: Wiley, 2001.

[26] F. Gustafsson, F. Gunnarsson, N. Bergman, U. Forssell, J. Jansson, R. Karlsson, and P.-J. Nordlund, "Particle filters for positioning, navigation, and tracking," *IEEE Trans. Signal Processing*, vol. 50, no. 2, pp. 425 –437, Feb. 2002.

[27] S. Ganeriwal, R. Kumar, and M. B. Srivastava, "Timing-sync protocol for sensor networks," in *Proc. ACM SenSys*, Los Angeles, CA, USA, 2003, pp. 138–149.

[28] J. Zheng and Y.-C. Wu, "Joint time synchronization and localization of an unknown node in wireless sensor networks," *IEEE Trans. Signal Processing*, vol. 58, no. 3, pp. 1309 –1320, Mar. 2010.

[29] B. Sundararaman, U. Buy, and A. Kshemkalyani, "Clock synchronization for wireless sensor networks: a survey," *Ad Hoc Networks*, vol. 3, no. 3, pp. 281 – 323, Jan. 2005.

[30] Y. Wang, G. Leus, and A.-J. van der Veen, "On digital receiver design for transmitted reference UWB," in *Proc. IEEE ICUWB*, Hannover, Germany, Sept. 2008, pp. 35–38.

[31] A. Schranzhofer, Y. Wang, and A.-J. van der Veen, "Acquisition for a transmitted reference UWB receiver," in *Proc. IEEE ICUWB*, Hannover, Germany, Sept. 2008, pp. 149–152.

[32] Y. Wang, G. Leus, and A.-J. van der Veen, "Digital receiver design for transmitted-reference ultra-wideband systems," *EURASIP J. Wireless Commun. Netw.*, vol. 2009, pp. 1–17, 2009.

[33] L. Yang and G. Giannakis, "Ultra-wideband communications: an idea whose time has come," *IEEE Signal Processing Mag.*, vol. 21, pp. 26–54, Nov. 2004.

[34] Z. Tian and G. Giannakis, "BER sensitivity to mistiming in ultra-wideband impulse radios –Part II: Performance in fading channels," *IEEE Trans. Signal Processing*, vol. 53, pp. 1897–1907, May 2005.

[35] R. Blazquez, P. Newaskar, and A. Chandrakasan, "Coarse acquisition for ultra wideband digital receivers," in *Proc. IEEE ICASSP*, vol. 4, Hong Kong, China, Apr. 2003, pp. 137–140.

[36] V. Lottici, A. D'Andrea, and U. Mengali, "Channel estimation for ultra-wideband communications," *IEEE J. Select. Areas Commun.*, vol. 20, pp. 1638–1645, Sept. 2002.

[37] S. Aedudodla, S. Vijayakunmaran, and T. Wong, "Timing acquisition in ultra-wideband communication systems," *IEEE Trans. Veh. Technol.*, vol. 54, pp. 1570–1583, Sept. 2005.

[38] Z. Tian and G. Giannakis, "A GLRT approach to data-aided timing acquisition in UWB radios –Part I: Algorithms," *IEEE Trans. Wireless Commun.*, vol. 4, pp. 2956–2967, Nov. 2005.

[39] J. Kusuma, I. Maravic, and M. Vetterli, "Sampling with finite rate of innovation: channel and timing estimation for UWB and GPS," in *Proc. IEEE ICC*, vol. 5, Anchorage, AK, May 2003, pp. 3540–3544.

[40] L. Yang and G. Giannakis, "Timing ultra-wideband signals with dirty templates," *IEEE Trans. Commun.*, vol. 53, pp. 1952–1963, Nov. 2005.

[41] I. Guvenc, Z. Sahinoglu, and P. Orlik, "TOA estimation for IR-UWB systems with different transceiver types," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 4, pp. 1876–1886, Jun. 2006.

[42] R. Hoctor and H. Tomlinson, "Delay-hopped transmitted-reference RF communications," in *Proc. IEEE Conf. UWB Syst. & Technol.*, Baltimore, MD, May 2002, pp. 265–269.

[43] M. Ho, V. Somayazulu, J. Foerster, and S. Roy, "A differential detector for an ultra-wideband communications system," in *Proc. IEEE VTC*, vol. 4, Birmingham, AL, USA, May 2002, pp. 1896–1900.

[44] Z. Tian and B. Sadler, "Weighted energy detection of ultra-wideband signals," in *Proc. IEEE SPAWC*, New York, NY, USA, Jun. 2005, pp. 1068–1072.

[45] Y. Vanderperren, G. Leus, and W. Dehaene, "A reconfigurable pulsed UWB receiver sampling below nyquist rate," in *Proc. IEEE ICUWB*, Hannover, Germany, Sept. 2008, pp. 145–148.

[46] J. Kim, S. Yang, and S. Y., "A two-step search scheme for rapid and reliable UWB signal acquisition in multipath channels," in *Proc. IEEE ICUWB*, Zurich, Switzerland, Sept. 2005, pp. 355–360.

[47] S. Gezici, Z. Sahinoglu, A. Molisch, H. Kobayashi, and H. Poor, "Two-step time of arrival estimation for pulse-based ultra-wideband systems," *EURASIP J. Advances in Signal Processing*, pp. 1–11, 2008.

[48] M. Casu and G. Durisi, "Implementation aspects of a transmitted-reference UWB receiver," *Wireless Communications and Mobile Computing*, vol. 5, pp. 537–549, 2005.

[49] S. Bagga, L. Zhang, W. Serdijin, J. Long, and E. Busking, "A quantized analog delay for an IR-UWB quadrature downconversion autocorrelation receiver," in *Proc. IEEE ICUWB*, Zurich, Switzerland, Sept. 2005, pp. 328–332.

[50] R. Qiu, H. Liu, and X. Shen, "Ultra-wideband for multiple access communications," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 80–87, Feb. 2005.

[51] R. Djapic, G. Leus, and A. van der Veen, "Blind synchronization in asynchronous multiuser UWB networks based on the transmit-reference scheme," in *Proc. IEEE Asilomar Conf. Signals, Systems and Computers*, vol. 2, Pacific Grove, CA, Nov. 2004, pp. 1506–1510.

[52] Q. Dang and A. van der Veen, "A decorrelating multiuser receiver for transmit-reference UWB systems," *IEEE J. Sel. Topics Signal Process.*, vol. 1, pp. 431–442, Oct. 2007.

[53] J. Choi and W. Stark, "Performance of ultra-wideband communications with suboptimal receivers in multipath channels," *IEEE J. Select. Areas Commun.*, vol. 20, no. 9, pp. 1754–1766, Dec. 2002.

[54] K. Witrisal, G. Leus, M. Pausini, and C. Krall, "Equivalent system model and equalization of differential impulse radio UWB systems," *IEEE J. Select. Areas Commun.*, vol. 23, pp. 1851–1862, Sept. 2005.

[55] T. Quek and M. Win, "Analysis of UWB transmitted-reference communication systems in dense multipath channels," *IEEE J. Select. Areas Commun.*, vol. 23, no. 9, pp. 1863–1874, Sept. 2005.

[56] T. Quek, M. Win, and D. Dardari, "Unified analysis of UWB transmitted-reference schemes in the presence of narrowband interference," *IEEE Trans. Wireless Commun.*, vol. 6, no. 6, pp. 2126–2139, Jun. 2007.

[57] S. Kay, *Fundamentals of Statistical Signal Processing: Vol. I – Estimation Theory*. Prentice-Hall, 1993.

[58] J. Foerster, "Channel modeling sub-committee report final," IEEE P802.15 Working Group for WPAN, Tech. Rep. IEEE P802.15-02/368r5-SG3a, 2002.

[59] H. Stark and J. Woods, *Probability, Random Processes, and Estimation Theory for Engineers*. Englewood Cliffs, NJ: Prentice-Hall, 1994.

[60] M. Simon, S. Hinedi, and W. Lindsey, *Digital Communication Techniques*. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[61] L. Yang and G. Giannakis, "Optimal pilot waveform assisted modulation for ultrawideband communications," *IEEE Trans. Wireless Commun.*, vol. 3, no. 4, pp. 1236–1249, Jul. 2004.

[62] Y. Wang, G. Leus, and A.-J. van der Veen, "Cramér-Rao bound for range estimation," in *Proc. IEEE ICASSP*, Taipei, Taiwan, Apr. 2009, pp. 3301–3304.

[63] J. Zhang, R. A. Kennedy, and T. D. Abhayapala, "Cramér-Rao lower bounds for the synchronization of UWB signals," *EURASIP J. Appl. Signal Processing*, vol. 3, pp. 426–438, 2005.

[64] R. Cardinali, L. De Nardis, M.-G. Di Benedetto, and P. Lombardo, "UWB ranging accuracy in high- and low-data-rate applications," *IEEE Trans. Microwave Theory Tech.*, vol. 54, pp. 1865–1875, Jun. 2006.

[65] Y. Qi and H. Kobayashi, "On relation among time delay and signal strength based geolocation methods," in *IEEE GLOBECOM*, vol. 7, San Francisco, CA, USA, Dec. 2003, pp. 4079–4083.

[66] A. Catovic and Z. Sahinoglu, "The Cramer-Rao bounds of hybrid TOA/RSS and TDOA/RSS location estimation schemes," *IEEE Commun. Lett.*, vol. 8, pp. 626–628, Oct. 2004.

[67] T. Rappaport, *Wireless communications principles and practice*. Prentice Hall, 1996.

[68] H. Van Trees, *Detection, estimation, and modulation theory*. New York: John Wiley & Sons, Inc., 1968.

[69] F. Gini and R. Reggiannini, "On the use of Cramer-Rao-like bounds in the presence of random nuisance parameters," *IEEE Trans. Commun.*, vol. 48, pp. 2120–2126, Dec. 2000.

[70] A. D'Andrea, U. Mengali, and R. Reggiannini, "The modified Cramer-Rao bound and its application to synchronization problems," *IEEE Trans. Commun.*, vol. 42, no. 234, pp. 1391–1399, Feb./Mar./Apr. 1994.

[71] R. Miller and C. Chang, "A modified Cramér-Rao bound and its applications," *IEEE Trans. Inform. Theory*, vol. 24, pp. 398–400, May 1978.

[72] B. Z. Bobrovsky, E. Mayer-Wolf, and M. Zakai, "Some classes of global Cramer-Rao bounds," *The Annals of Statistics*, vol. 15, no. 4, pp. 1421–1438, Dec. 1987.

[73] I. Reuven and H. Messer, "A Barankin-type lower bound on the estimation error of a hybrid parameter vector," *IEEE Trans. Inform. Theory*, vol. 43, pp. 1084–1093, May 1997.

[74] J.-Y. Lee and R. Scholtz, "Ranging in a dense multipath environment using an UWB radio link," *IEEE J. Select. Areas Commun.*, vol. 20, no. 9, pp. 1677–1683, Dec. 2002.

[75] IEEE Working Group 802.15.4, "Part 15.4: Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate wireless personal area networks (WPANs)," Tech. Rep., 2007.

[76] Z. Sahinoglu and S. Gezici, "Ranging in the IEEE 802.15.4a standard," in *Proc. IEEE WAMICON*, Clearwater, FL, USA, Dec. 2006, pp. 1–5.

[77] D. Dardari, A. Conti, U. Ferner, A. Giorgetti, and M. Z. Win, "Ranging with ultrawide bandwidth signals in multipath environments," *Proc. IEEE*, vol. 97, no. 2, pp. 404–426, Feb. 2009.

[78] J. Zhang, P. V. Orlik, Z. Sahinoglu, A. F. Molisch, and P. Kinney, "UWB systems for wireless sensor networks," *Proc. IEEE*, vol. 97, no. 2, pp. 313–331, Feb. 2009.

[79] I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Commun. Mag.*, vol. 40, no. 8, pp. 102 – 114, Aug. 2002.

[80] C.-Y. Chong and S. Kumar, "Sensor networks: evolution, opportunities, and challenges," *Proc. IEEE*, vol. 91, no. 8, pp. 1247–1256, Aug. 2003.

[81] M. G. Verdone R., Dardari D. and C. A., *Wireless Sensor and Actuator Networks: Technologies, Analysis and Design.* Amsterdam, the Netherlands: Elsevier, 2008.

[82] I. Guvenc and Z. Sahinoglu, "Threshold-based TOA estimation for impulse radio UWB systems," in *Proc. IEEE ICUWB*, Zurich, Switzerland, Sept. 2005, pp. 420–425.

[83] C. Falsi, D. Dardari, L. Mucchi, and M. Win, "Time of arrival estimation for UWB localizers in realistic environments," *EURASIP J. Appl. Signal Processing*, pp. 1–13, 2006.

[84] D. Dardari, C.-C. Chong, and M. Win, "Threshold-based time-of-arrival estimators in UWB dense multipath channels," *IEEE Trans. Commun.*, vol. 56, no. 8, pp. 1366–1378, Aug. 2008.

[85] A. D'Amico, U. Mengali, and L. Taponecco, "Energy-based TOA estimation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 838–847, Mar. 2008.

[86] R. Zhang and X. Dong, "A new time of arrival estimation method using UWB dual pulse signals," *IEEE Trans. Wireless Commun.*, vol. 7, no. 6, pp. 2057–2062, Jun. 2008.

[87] I. Guvenc, S. Gezici, and Z. Sahinoglu, "Ultra-wideband range estimation: theoretical limits and practical algorithms," in *Proc. IEEE ICUWB*, vol. 3, Hannover, Germany, Sept. 2008, pp. 93–96.

[88] I. Guvenc, Z. Sahinoglu, A. Molisch, and R. Orlik, "Non-coherent TOA estimation in IR-UWB systems with different signal waveforms," in *Proc. IEEE BroadNets*, Boston, MA, USA, Oct. 2005, pp. 1168–1174.

[89] Z. Irahhauten, A. Yarovoy, G. Janssen, H. Nikookar, and L. Ligthart, "Suppression of noise and narrowband interference in UWB indoor channel measurements," in *Proc. IEEE ICUWB*, Zurich, Switzerland, Sept. 2005, pp. 108–112.

[90] B. Zhen, H.-B. Li, and R. Kohno, "Clock management in ultra-wideband ranging," in *Proc. IST Mobile and Wireless Commun. Summit*, Budapest, Hungary, Jul. 2007, pp. 1–5.

[91] A. Wellig and Y. Qiu, "Trellis-based maximum-likelihood crystal drift estimator for ranging applications in UWB-LDR," in *Proc. IEEE ICUWB*, Waltham, MA, USA, Sept. 2006, pp. 539–544.

[92] A. Wellig, "Method and device for estimation the relative drift between two clocks, in particular for ranging applications in UWB-LDR technology," U.S. Patent 2008/0 069 260A1, Mar. 20, 2008.

[93] L. Huang and C. C. Ko, "Performance of maximum-likelihood channel estimator for UWB communications," *IEEE Commun. Lett.*, vol. 8, no. 6, pp. 356 – 358, Jun. 2004.

[94] Y. Chen and N. Beaulieu, "CRLBs for NDA ML estimation of UWB channels," *IEEE Commun. Lett.*, vol. 9, no. 8, pp. 709 – 711, Aug. 2005.

[95] W. Black and D. Hodges, "Time interleaved converter arrays," *IEEE J. Solid-State Circuits*, vol. 15, pp. 1022–1029, Dec. 1980.

[96] B. Razavi, "Design of sample-and-hold amplifiers for high-speed low-voltage A/D converters," in *Proc. IEEE CICC*, Santa Clara, CA, USA, May 1997, pp. 59–66.

[97] M. Shinagawa, Y. Akazawa, and T. Wakimoto, "Jitter analysis of high-speed sampling systems," *IEEE J. Solid-State Circuits*, vol. 25, no. 1, pp. 220–224, Feb. 1990.

[98] K. Poulton, J. Corcoran, and T. Hornak, "A 1-GHz 6-bit ADC system," *IEEE J. Solid-State Circuits*, vol. 22, no. 6, pp. 962–970, Dec. 1987.

[99] D. Fu, K. Dyer, S. Lewis, and P. Hurst, "A digital background calibration technique for time-interleaved analog-to-digital converters," *IEEE J. Solid-State Circuits*, vol. 33, no. 12, pp. 1904–1911, Dec. 1998.

[100] S. Kay, *Fundamentals of Statistical Signal Processing: Vol. II – Detection Theory*. Prentice-Hall, 1993.

[101] S. Wu, Q. Zhang, R. Fan, and N. Zhang, "Match-filtering based TOA estimation for IR-UWB ranging systems," in *Proc. IEEE IWCMC*, Crete Island, Greece, Aug. 2008, pp. 1099–1105.

[102] A. Molisch, K. Balakrishnan, C. Chong, S. Emami, A. Fort, J. Karedal, J. Kunisch, H. Schantz, U. Schuster, and K. Siwiak, "IEEE 802.15.4a channel model-final report," Tech. Rep., 2005.

[103] B. Sadler, L. Huang, and Z. Xu, "Ziv-Zakai time delay estimation bound for ultra-wideband signals," in *Porc. IEEE ICASSP*, vol. 3, Apr. 2007, pp. 549 –552.

[104] G. Golub and C. van Loan, *Matrix Computations*. The Johns Hopkins Univ. Press, 1996.

[105] A. Amar and G. Leus, "A reference-free time difference of arrival source localization using a passive sensor array," in *IEEE SAM*, Jerusalem, Israel, Oct. 2010, pp. 157 –160.

[106] W. S. Torgerson, "Multidimensional scaling: I. theory and method." *Psychometrika*, vol. 17, pp. 401–419, 1952.

[107] J. A. Costa, N. Patwari, and A. Hero III, "Distributed weighted multidimensional scaling for node localization in sensor networks," *ACM Trans. on Sensor Networks*, vol. 2, pp. 1–26, Feb. 2006.

[108] H. W. Weim, Q. Wan, Z. X. Chen, and S. F. Ye, "A novel weighted multidimensional analysis for time of arrival based mobile location," *IEEE Trans. Signal Processing*, vol. 56, no. 7, pp. 3018–3022, Jul. 2008.

[109] K. W. Cheung and H. So, "A multidimensional scaling framework for mobile location using time of arrival measurements," *IEEE Trans. Signal Processing*, vol. 53, no. 2, pp. 460–470, Feb. 2005.

[110] Y. Shang and W. Ruml, "Improved MDS-based localization," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar. 2004, pp. 2640–2651.

[111] X. Ji and H. Zha, "Sensor positioning in wireless ad hoc sensor networks using multidimensional scaling," in *Proc. IEEE INFOCOM*, Hong Kong, China, Mar. 2004, pp. 2652–2661.

[112] I. Borg and P. Groenen, *Modern Multidimensional Scaling Theory and Applications*. Berlin, Germany: Springer, 1997.

[113] P. Drineas, M. Magdon-Ismail, G. Pandurangan, R. Virrankoski, and A. Savvides, "Distance matrix reconstruction from incomplete distance information for sensor network localization," in *Proc. IEEE SECON*, vol. 2, Reston, VA, USA, Sep. 2006, pp. 536–544.

[114] J. C. Platt, "Fastmap, metricmap, and landmark MDS are all Nyström algorithm," in *Proc. AISTATS*, Bridgetown, Barbados, Jan. 2005, pp. 261–268.

[115] Y. Wang and G. Leus, "Reference-free time-based localization for an asynchronous target," *EURASIP J. Advances in Signal Processing*, Sept. 2011, accepted.

[116] Y. Wang, G. Leus, and X. Ma, "Tracking a mobile node by asynchronous networks," in *Proc. IEEE SPAWC*, San Francisco, CA, USA, Jun. 2011, pp. 1–5.

[117] J. Yan, *Algorithms for indoor positioning systems using ultra-wideband signals*, Delft University of Technology, 2010, PhD Thesis.

[118] Z. Li, W. Trappe, Y. Zhang, and B. Nath, "Robust statistical methods for securing wireless localization in sensor networks," in *Proc. IPSN*, Los Angeles, CA, USA, Apr. 2005, pp. 91–98.

[119] S. Venkatesh and R. Buehrer, "A linear programming approach to NLOS error mitigation in sensor networks," in *Proc. IPSN*, Nashville, TN, USA, Apr. 2006, pp. 301 –308.

[120] J. Smith and J. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 12, pp. 1661–1669, Dec. 1987.

[121] B. Friedlander, "A passive localization algorithm and its accurancy analysis," *IEEE J. Oceanic Eng.*, vol. 12, no. 1, pp. 234–245, Jan. 1987.

[122] Y. Chan and K. Ho, "A simple and efficient estimator for hyperbolic location," *IEEE Trans. Signal Processing*, vol. 42, no. 8, pp. 1905–1915, Aug. 1994.

[123] P. Stoica and J. Li, "Lecture notes - source localization from range-difference measurements," *IEEE Signal Processing Mag.*, vol. 23, no. 6, pp. 63–66, Nov. 2006.

[124] F. Chan and H. So, "Efficient weighted multidimensional scaling for wireless sensor network localization," *IEEE Trans. Signal Processing*, vol. 57, no. 11, pp. 4548 –4553, Nov. 2009.

[125] Y. Bresler and A. Macovski, "Exact maximum likelihood parameter estimation of superimposed exponential signals in noise," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 34, no. 5, pp. 1081 – 1089, Oct. 1986.

[126] A. Beck, P. Stoica, and J. Li, "Exact and approximate solutions of source localization problems," *IEEE Trans. Signal Processing*, vol. 56, no. 5, pp. 1770 –1778, May 2008.

[127] S. Zhu and Z. Ding, "Joint synchronization and localization using TOAs: A linearization based WLS solution," *IEEE J. Select. Areas Commun.*, vol. 28, no. 7, pp. 1017 – 1025, Sept. 2010.

[128] S. Bancroft, "An algebraic solution of the GPS equations," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 21, no. 1, pp. 56 –59, Jan. 1985.

[129] H. Schau and A. Robinson, "Passive source localization employing intersecting spherical surfaces from time-of-arrival differences," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 35, no. 8, pp. 1223 – 1225, Aug. 1987.

[130] S. Barnet, *Matrices, Methods and Applications*. Oxford Applied Mathematics and Computin Science Series, Clarendon Press, 1990.

[131] D.-H. Shin and T.-K. Sung, "Comparisons of error characteristics between TOA and TDOA positioning," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 38, no. 1, pp. 307 –311, 2002.

[132] H. C. So, Y. T. Chan, and F. Chan, "Closed-form formulae for time-difference-of-arrival estimation," *IEEE Trans. Signal Processing*, vol. 56, no. 6, pp. 2614 –2620, Jun. 2008.

[133] M. Gillette and H. Silverman, "A linear closed-form algorithm for source localization from time-differences of arrival," *IEEE Signal Processing Lett.*, vol. 15, pp. 1 –4, 2008.

[134] J. Smith and J. Abel, "The spherical interpolation method of source localization," *IEEE J. Oceanic Eng.*, vol. 12, no. 1, pp. 246 – 252, Jan. 1987.

[135] Y. Wang, X. Ma, and G. Leus, "An UWB ranging-based localization strategy with internal attack immunity," in *Proc. IEEE ICUWB*, Nanjing, China, Sept. 2010, pp. 1–4.

[136] Y. Wang, X. Ma, , and G. Leus, "Robust time-based localization for asynchronous networks," *IEEE Trans. Signal Processing*, vol. 59, no. 9, pp. 4397–4410, Sept. 2011.

[137] Y. Wang, G. Leus, and X. Ma, "Time-based localization for asynchronous wireless sensor networks," in *Proc. IEEE ICASSP*, Prague, Czech Republic, May 2011, pp. 3284–3287.

[138] Y. Wang, X. Ma, and G. Leus, "Robust time-based localization for asynchronous networks with clock offsets," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, Nov. 2011, accepted.

[139] O. Simeone, U. Spagnolini, Y. Bar-Ness, and S. Strogatz, "Distributed synchronization in wireless networks," *IEEE Signal Processing Mag.*, vol. 25, no. 5, pp. 81 –97, Sept. 2008.

[140] I. Rhee, J. Lee, J. Kim, E. Serpedin, and Y.-C. Wu, "Clock synchronization in wireless sensor networks: an overview," *Sensors*, vol. 9, pp. 56–85, 2009.

[141] K. Römer and F. Mattern, "Towards a unified view on space and time in sensor networks," *Comput. Commun.*, vol. 28, no. 13, pp. 1484 – 1497, Aug. 2005.

[142] J. Elson, L. Girod, and D. Estrin, "Fine-grained network time synchronization using reference broadcasts," *SIGOPS Oper. Syst. Rev.*, vol. 36, pp. 147–163, Dec. 2002.

[143] M. Maróti, B. Kusy, G. Simon, and A. Lédeczi, "The flooding time synchronization protocol," in *Proc. ACM SenSys*, Baltimore, MD, USA, Nov. 2004, pp. 39–49.

[144] D. Jeske, "On maximum-likelihood estimation of clock offset," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 53 – 54, Jan. 2005.

[145] K.-L. Noh, Q. Chaudhari, E. Serpedin, and B. Suter, "Novel clock phase offset and skew estimation using two-way timing message exchanges for wireless sensor networks," *IEEE Trans. Commun.*, vol. 55, no. 4, pp. 766 –777, Apr. 2007.

[146] Q. Chaudhari, E. Serpedin, and K. Qaraqe, "On maximum likelihood estimation of clock offset and skew in networks with exponential delays," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1685 –1697, Apr. 2008.

[147] X. Cheng, A. Thaeler, G. Xue, and D. Chen, "TPS: a time-based positioning scheme for outdoor wireless sensor networks," in *Proc. IEEE INFOCOM*, vol. 4, Hong Kong, China, Mar. 2004, pp. 2685 – 2696.

[148] H. de Oliveira, E. Nakamura, A. Loureiro, and A. Boukerche, "Localization in time and space for sensor networks," in *Proc. AINA*, Niagara Falls, Ontario, Canada, May 2007, pp. 539 –546.

[149] B. Denis, J.-B. Pierrot, and C. Abou-Rjeily, "Joint distributed synchronization and positioning in UWB ad hoc networks using TOA," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 4, pp. 1896 – 1911, Jun. 2006.

[150] C. Yan and H. Fan, "Asynchronous self-localization of sensor networks with large clock drift," in *Proc. IEEE MobiQuitous*, Philadelphia, PA, USA, Aug. 2007, pp. 1–8.

[151] K. Yu, Y. Guo, and M. Hedley, "TOA-based distributed localisation with unknown internal delays and clock frequency offsets in wireless sensor networks," *IET Signal Processing*, vol. 3, no. 2, pp. 106 –118, 2009.

[152] Y. Zhou, C. L. Law, Y. L. Guan, and F. Chin, "Indoor elliptical localization based on asynchronous uwb range measurement," *IEEE Trans. Instrum. Meas.*, vol. 60, no. 1, pp. 248–257, Jan. 2011.

[153] L. Lazos and R. Poovendran, "SeRLoc: secure range-independent localization for wireless sensor networks," in *Proc. ACM WiSe*, Philadelphia, PA, USA, Oct. 2004, pp. 21–30.

[154] S. Capkun and J.-P. Hubaux, "Secure positioning of wireless devices with application to sensor networks," in *Proc. IEEE INFOCOM*, vol. 3, Miami, FL, USA, Mar. 2005, pp. 1917 – 1928.

[155] Y. Chen, W. Trappe, and R. P. Matrin, "Attack detection in wireless localization," in *Proc. IEEE INFOCOM*, Anchorage, AK, USA, May 2007, pp. 1964–1972.

[156] D. Liu, P. Ning, A. Liu, C. Wang, and W. K. Du, "Attack-resistant location estimation in wireless sensor networks," *ACM Trans. Inf. Syst. Secur.*, vol. 11, no. 4, pp. 1–39, Jul. 2008.

[157] B. Denis and N. Daniele, "NLOS ranging error mitigation in a distributed positioning algorithm for indoor UWB ad-hoc networks," in *Proc. IEEE IWWAN*, Oulu, Finland, May 2004, pp. 356 – 360.

[158] H. Kopetz and W. Ochsenreiter, "Clock synchronization in distributed real-time systems," *IEEE Trans. Commun.*, vol. C-36, no. 8, pp. 933 –940, Aug. 1987.

[159] Z. Sahinoglu and S. Gezici, "Enhanced position estimation via node cooperation," in *Proc. IEEE ICC*, Cape Town, South Africa, May 2010, pp. 1–6.

[160] G. J. Pottie and W. J. Kaiser, "Wireless integrated network sensors," *Commun. of the ACM*, vol. 43, no. 5, pp. 51–58, May 2000.

[161] K. Yu, J.-P. Montillet, A. Rabbachin, P. Cheong, and I. Oppermann, "UWB location and tracking for wireless embedded networks," *Signal Processing*, vol. 86, no. 9, pp. 2153 – 2171, 2006.

[162] T. Li, A. Ekpenyong, and Y.-F. Huang, "Source localization and tracking using distributed asynchronous sensors," *IEEE Trans. Signal Processing*, vol. 54, no. 10, pp. 3991 –4003, 2006.

[163] M. Vemula, J. Miguez, and A. Artes-Rodriguez, "A sequential Monte Carlo method for target tracking in an asynchronous wireless sensor network," in *Proc. WPNC*, Hannover, Germany, Mar. 2007, pp. 49 –54.

[164] S. Julier and J. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401 – 422, Mar. 2004.

[165] B. R. Hamilton, X. Ma, Q. Zhao, and J. Xu, "ACES: adaptive clock estimation and synchronization using Kalman filtering," in *Proc. ACM MobiCom*, New York, NY, USA, 2008, pp. 152–162.

[166] R. Doraiswami, "A novel Kalman filter-based navigation using beacons," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 830 –840, Apr. 1996.

[167] R. Moose, H. Vanlandingham, and D. Mccabe, "Modeling and estimation for tracking maneuvering targets," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 15, no. 3, pp. 448 –456, May 1979.

[168] Y. Chan, A. Hu, and J. Plant, "A Kalman filter based tracking scheme with input estimation," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 15, no. 2, pp. 237 –244, Mar. 1979.

[169] P. Biswas and Y. Ye, "Semidefinite programming for ad hoc wireless sensor network localization," in *Proc. IPSN*, Berkeley, CA, USA, Apr. 2004, pp. 46–54.

[170] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Mag.*, vol. 27, no. 3, pp. 20 –34, May 2010.

[171] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," in *Proc. IPSN*, Los Angeles, CA, USA, Apr. 2005, pp. 63 – 70.

[172] I. Schizas, A. Ribeiro, and G. Giannakis, "Consensus in ad hoc wsns with noisy links – Part I: Distributed estimation of deterministic signals," *IEEE Trans. Signal Processing*, vol. 56, no. 1, pp. 350 –364, Jan. 2008.

[173] I. Schizas, G. Giannakis, S. Roumeliotis, and A. Ribeiro, "Consensus in ad hoc wsns with noisy links – Part II: Distributed estimation and smoothing of random signals," *IEEE Trans. Signal Processing*, vol. 56, no. 4, pp. 1650 –1666, Apr. 2008.

[174] M. Rabbat, R. Nowak, and J. Bucklew, "Generalized consensus computation in networked systems with erasure links," in *Proc. IEEE SPAWC*, New York, NY, USA, Jun. 2005, pp. 1088 – 1092.

[175] S. Srirangarajan, A. Tewfik, and Z.-Q. Luo, "Distributed sensor network localization using SOCP relaxation," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4886 –4895, Dec. 2008.

[176] P. J. Huber, "Robust regression: Asymptotics, conjectures and monte carlo," *The Annals of Statistics*, vol. 1, no. 5, pp. 799–821, Sep. 1973.

[177] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. John Wiley & Sons, 1987.

[178] J.-J. Fuchs, "An inverse problem approach to robust regression," in *Proc. IEEE ICASSP*, vol. 4, Phoenix, AZ, USA, Mar. 1999, pp. 1809 –1812.

[179] V. Kekatos and G. Giannakis, "From sparse signals to sparse residuals for robust sensing," *IEEE Trans. Signal Processing*, vol. 59, no. 7, pp. 3355 –3368, july 2011.

[180] L. Xie, Y. C. Soh, and C. de Souza, "Robust Kalman filtering for uncertain discrete-time systems," *IEEE Trans. Automat. Contr.*, vol. 39, no. 6, pp. 1310 –1314, Jun. 1994.

[181] I. R. Petersen, *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*. Birkhauser Boston, 1999.

[182] A. Sayed, "A framework for state-space estimation with uncertain models," *IEEE Trans. Automat. Contr.*, vol. 46, no. 7, pp. 998 –1013, Jul. 2001.

[183] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. IEEE CDC-ECC*, Seville, Spain, Dec. 2005, pp. 8179 – 8184.

[184] ——, "Distributed kalman filtering for sensor networks," in *Proc. IEEE CDC-ECC*, New Orleans, LA, USA, Dec. 2007, pp. 5492 –5498.

Samenvatting

De grote vraag naar plaatsbewuste draadloze sensornetwerken motiveert het onderzoek in dit proefschrift. De unieke eigenschappen van draadloze sensornetwerken geven tal van uitdagingen op het gebied van plaatsbepaling en communicatie. In dit proefschrift behandelen we een aantal belangrijke uitdagingen en bieden betaalbare oplossingen.

Impuls-radio ultra breedband (IR-UWB) wordt gebruikt als de onderliggende technologie voor zowel plaatsbepaling als communicatie vanwege de specifieke voordelen die het biedt voor nauwkeurige afstandsbepaling en betrouwbare communicatie. De volgende aspecten worden behandeld in dit proefschrift.

- *Meegestuurde-referentie (TR) UWB communicatiesystemen:* IR-UWB demodulatie in het digitale domein vraagt meestal om zeer hoge bemonsteringfrequenties. Het TR-UWB systeem zorgt voor sub-Nyquist bemonstering door middel van het correleren van de ontvangen pulsreeks met een vertraagde versie in het analoge tijddomein. Zodanig wordt de vrijwel onmogelijke Nyquist bemonstering vermeden, worden de strenge eisen aan de synchronisatie verlaagd, en zijn enkel de uiteindelijke kanaalcoëfficiënten nodig. Het gebruikte datamodel wordt uitgebreid om ook de verschillende vormen van interferentie te omvatten, en vervolgens wordt een complete ontvanger voorgesteld, waaronder signaaldetectie, kanaalschatting, synchronisatie en egalisatie.

- *Theoretische grenzen aan de afstandsbepaling en praktische afstandsbepalingstechnieken gebaseerd op IR-UWB:* We onderzoeken de theoretische nauwkeurigheid van de afstandsbepaling met een nieuwe methode, die de lokatie-informatie benut die aanwezig is in zowel de amplitude als de tijdsvertraging van het ontvangen signaal. Dit wordt niet enkel onderzocht voor gedempte kanalen met toegevoegde witte Gaussische ruis (AWGN), maar ook voor een AWGN kanaal met zowel demping als schaduweffecten. Verder is een praktische afstandsbepalingsmethode gebaseerd op een schatting van de tijd van aankomst (TOA) van het UWB-IR signaal ontwikkeld. Stroboscopis-

211

che bemonstering wordt gebruikt om transmissie-efficiëntie uit te wisselen voor een lagere bemonsterfrequentie, terwijl dezelfde afstandsresolutie als bij Nyquist bemonstering wordt bereikt. Doordat een lange uitbreiding voorafgaand aan het signaal vereist is voor stroboscopisch bemonsteren, is de klok drift (een cumulatief effect dat na verloop van tijd veroorzaakt wordt door de relatieve klokverschuivingen tussen twee klokken) een van de belangrijkste foutbronnen in de TOA schattingstechniek. Daarom worden schattingsmethoden voor de TOA inclusief klokdrift-kalibratie onderzocht om de invloed van de drift drastisch te verminderen.

- *Diverse plaatsbepalings- en volgmethoden:*

  – *Uitgebreide multi-dimensionale schaling (MDS):* Omdat het klassieke MDS algoritme niet kan worden toegepast op algemene netwerken met ontbrekende verbindingen, breiden we het MDS-algoritme uit om om te kunnen gaan met een speciaal soort netwerk met specifieke ontbrekende verbindingen. Ons doel is om, gegeven paarsgewijze afstandsmetingen, een gemeenschappelijke schatting te bepalen van de posities van alle knooppunten, afgezien van een een translatie, rotatie en reflectie.

  – *Referentie-vrije tijdgebaseerde plaatsbepaling:* Uit de literatuur zijn diverse kleinste kwadraten (LS) schatters met lage complexiteit bekend, gebaseerd op metingen van de aankomsttijd (TOA) of verschillen in aankomsttijd (TDOA), met het doel een object te lokaliseren met de hulp van ankers (knooppunten met bekende posities). Hierbij moet een anker aangewezen worden als referentiepunt, om hierdoor additionele parameters te verwijderen en de sterke eisen aan synchronisatie af te zwakken. Er is een probleem als hierbij een slechte keuze gemaakt wordt.

  We ontwikkelen een algoritme voor referentie-vrije lokatie-schatting, gebaseerd op TOA metingen, om de afhankelijkheid van de referentiekeuze te ontlopen. Verder generaliseren we bestaande referentiegebaseerde gesloten-vorm oplossingen die gebruik maken van TOA of TDOA metingen, en werpen een nieuw licht op hun relaties om hiermee een aantal verwarringen die nog bestaan in de recente literatuur te verduidelijken.

  – *Robuuste tijd-gebaseerde plaatsbepaling:* Tijd-gebaseerde plaatsbepalingsalgoritmes staan in de belangstelling vanwege hun hoge nauwkeurigheid en potentieel lage kosten voor WSNs. Echter, deze vorm van lokalisatie is nauw verbonden met kloksynchronisatie. Zo is de betrouwbaarheid van de tijdmetingen een belangrijk maar moeilijk aspect om rekening mee te houden. Zonder gebruikmaking van de tijdmetingen door het doelobject, stellen wij een nieuw afstandsbepalingsprotocol voor, genaamd asymmetrische trip afstandsbepaling (ATR), welke leidt tot lokalisatiemethoden die van nature immuun zijn voor

aanvallen opgezet door een gecompromitteerd doelobject. Robuuste lokalisatiestrategieën op basis van TOA metingen met behulp van het ATR-protocol worden voorgesteld, om hiermee een doelobject te lokaliseren met de hulp van ankers voor asynchrone netwerken.

– *Kalman volgers:* Vanwege de niet-lineariteit van het plaatsbepalings-probleem wordt een Kalman filter (KF) meestal vervangen door een uitgebreid KF (EKF) voor het volgen van een mobiel doel. Echter, de modelfouten die inherent vervat zijn in het EKF degraderen de prestatie hiervan. Daarom maken we opnieuw gebruik van het ATR-protocol, ontwikkelen we exacte lineariseringen, en bereiken we een KF gebaseerd op een lineair model voor de metingen, om een mobiel opject te volgen met de hulp van vaste ankers.

Acknowledgements

Let me first thank you all who inspired, encouraged and helped me during the past years, and thank my committee members for their time and valuable comments.

I offer my sincere gratitude to my promotor Prof. Alle-Jan van der Veen, who shows me the beauty and elegance of signal processing, encourages me to pursue my PhD, and has stimulating discussions with me.

Many thanks to my supervisor Dr. Geert Leus. When I come up with new ideas and questions, his instantaneous feedback always hits the hardcore. When I feel no way to go further, he sheds light on possible directions. When I just look from a single angle, he reminds me with other perspectives. I cannot accomplish this thesis without his smart guidance, stimulating discussions, abundant patience, generous kindness and also must-needed criticism.

Special thanks go to Prof. Xiaoli Ma. I am very lucky to come across with her. The half-year visit in Xiaoli's group not only resulted in a very fruitful research, but also opened a new door for me. She always encourages me to face new challenges. I have learned a lot from her technically and non-technically. She also shows me an example how to maintain a great balance between work and family.

I also want to thank Dr. Rene van Leuken, who is my master supervisor. He first introduced me to Alle-Jan, and motivated me to change the field from digital circuits design to signal processing.

During my stay in the CAS group, I am grateful for insights, inspirations, discussions and suggestions from our visiting members: Profs. Hakan Deliç, Zhi Tian, Amir Leshem and Urbashi Mitra, Dr. Alon Amar, Dr. Luca Rugini, Andreas Schranzhofer, and Christoph Krall. The fruitful collaboration results are collected in this thesis. I also enjoyed a lot to work in 3UB with Dr. Zoubir Irahhauten, Koen Ruskamp and Alexandro Mancusi. Our team work turns a concept design into a real prototype system.

It has been a great pleasure working with so many nice colleagues: Minaksie, Laura, Rosario, Antoon, Wim, Huib, Yu, Claude, Tang, Kun, Hieu, Bas, Tao Wang, Vijay, Seyran, Millad, Sina, Sharzard, Dony, Sundeep, Hadi, Hamid, Raj,

About the Author

Yiyin Wang was born in Chengdu, China, in 1980. She received the B.S. degree in electrical engineering from Fudan University, Shanghai, China, in 2002, and the M.S. degree (cum laude) in microelectronics from the Delft University of Technology, The Netherlands, and Fudan University, Shanghai, China, in 2005, respectively. From 2005 to 2006, she was a Research Assistant in ASIC & System State Key Lab, Fudan University. From 2006 to 2007, she was a Research Assistant in the Circuits and Systems group (CAS) at the Faculty of Electrical Engineering, Mathematics and Computer Science (EEMCS) of the Delft University of Technology, The Netherlands. Since Nov. 2007, she has been working towards the Ph.D. degree in the same group. From February 2010 till July 2010, she was a Visiting Student at Georgia Institute of Technology, Atlanta, USA. Her research interests lie in the area of signal processing for localization and wireless communications employing tools from linear algebra and statistical signal processing.

Publications

Journal Papers:

[1] Y. Wang and G. Leus, "Reference-free Time-based Localization for an Asynchronous Target", *EURASIP Journal on Advances in Signal Processing*, accepted, Sept. 2011.

[2] Y. Wang, X. Ma and G. Leus, "Robust Time-based Localization for Asynchronous Networks", *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp.4397-4410, Sept. 2011.

[3] A. Amar, Y. Wang and G. Leus, "Extending the Classical Multidimensional Scaling Algorithm Given Partial Pairwise Distance Measurements", *IEEE Signal Processing Letters*, vol. 17, no. 5, pp. 473-476, May 2010.

[4] Y. Wang, G. Leus and A.-J. van der Veen, "Digital Receiver Design for Transmitted-Reference Ultra-Wideband Systems", *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 1-17, Feb. 2009, articleID 315264, doi:10.1155/2009/315264.

Conference Papers:

[5] Y. Wang, X. Ma and G. Leus, "Robust Time-Based Localization for Asynchronous Networks with Clock Offsets", in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, USA, accepted, Nov. 2011.

[6] Y. Wang, G. Leus and X. Ma, "Tracking a Mobile Node by Asynchronous Networks", in *Proc. IEEE SPAWC*, San Francisco, CA, USA, Jun. 2011.

[7] Y. Wang, G. Leus and X. Ma, "Time-based Localization for Asynchronous Wireless Sensor Networks", in *Proc. IEEE ICASSP*, Prague, Czech Republic, pp. 3284-3287, May 2011.

[8] Y. Wang, X. Ma and G. Leus, "A UWB Ranging-based Localization Strategy with Internal Attack Immunity", in *Proc. IEEE ICUWB*, Nanjing, China, pp. 1-4, Sept. 2010.

[9] Y. Wang, G. Leus and H. Delic, "TOA Estimation Using UWB with Low Sampling Rate and Clock Drift Calibration", in *Proc. IEEE ICUWB*, Vancouver, Canada, pp. 612-617, Sept. 2009.

[10] Y. Wang, G. Leus and A.J. van der Veen, "Cramer-Rao Bound for Range Estimation", in *Proc. IEEE ICASSP*, Taipei, Taiwan, pp. 3301-3304, Apr. 2009.

[11] Y. Wang, G. Leus and A.J. van der Veen, "On Digital Receiver Design for Transmitted Reference Ultra Wideband", in *Proc. IEEE ICUWB*, Hannover, Germany, pp. 35-38, Sept. 2008.

[12] A. Schranzhofer, Y. Wang and A.J. van der Veen, "Acquisition for a Transmitted Reference UWB Receiver", in *Proc. IEEE ICUWB*, Hannover, Germany, pp. 149-152, Sept. 2008.

[13] Y. Wang, A. Schranzhofer, R. van Leuken and A.J. van der Veen, "Hardware Platform for Delay Hopped Transmitted Reference UWB Communication System Prototype Development", in *Proc. IEEE ProRISC*, Veldhoven, The Netherlands, pp. 272-275, Nov. 2007.

[14] Y. Wang, A.J. van der Veen and R. van Leuken, "Design of a Practical Scheme for Ultra Wideband Communication", in *Proc. IEEE ISCAS*, Kos, Greek, pp. 2585-2588, May 2006.