Radio Frequency Ranging for Precise Indoor Localization

DISSERTATION
zur Erlangung des akademischen Grades

Doktor-Ingenieur
(Dr.-Ing.)
im Fach Informatik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von
Dipl.-Ing. Vladica Sark

Präsidentin der der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Mathematisch-Naturwissenschaftlichen Fakultät:
Prof. Dr. Elmar Kulke

Gutachter/innen:

1. Prof. Dr. Eckhard Graß	Humboldt-Universität zu Berlin
2. Prof. Dr. Mesut Günes	Universität Magdeburg
3. Prof. Dr. Verena Hafner	Humboldt-Universität zu Berlin

Tag der mündlichen Prüfung: 17. Juli 2017

To my beloved wife Biljana
and our dear children Johan and Dario.

Acknowledgements

Preparing this thesis would have not been possible without support of many people.

I would like to thank my supervisor Prof. Dr. Eckhard Grass, for his support and for many discussions which brought new ideas. I would also like to thank my colleague Jesús Gutiérrez Terán, who provided valuable comments and suggestions which significantly improved the quality of this work.

I would also like to express gratitude to my colleagues at IHP for the great support, especially to the members of the broadband communications group for the many discussions which helped refining my ideas.

I would also like to thank my colleagues Frank Winkler and Felix Wermke from the Humboldt University of Berlin for their valuable opinions on the work in this thesis.

This research was partially supported by the German Research Foundation (Deutsche Forschungsgemeinschaft DFG) funded graduate college METRIK. I am very thankful for this opportunity.

Finally, this work would have not been possible without the support of my wife and without the happiness our children Johan and Dario brought in our lives. Thank you.

Abstract

In the last couple of decades the Global Navigation Satellite Systems (GNSS) have become a very important part of our everyday life. A huge number of applications offer location based services and navigation functions which rely on these systems. Nevertheless, the offered localization services are not available indoors and their performance is significantly affected in urban areas. Therefore, in the recent years, a large number of wireless indoor localization systems are being actively investigated and developed. Widely known localization methods based on time of flight (ToF), angle of arrival (AoA) or received signal strength are commonly used in indoor positioning systems. On the one hand, the localization precision and accuracy offered is required to be significantly higher compared to the global and large area positioning systems. On the other hand, their complexity and cost must be much lower, since they have notably smaller coverage.

The main focus of this work is on improving precision and accuracy of indoor localization systems, as well as on the implementation and integration of localization functionality in wireless data transmission systems. Two approaches for improving the localization precision and accuracy of ToF based methods are proposed. The first approach, referred to as modified equivalent time sampling (METS) is used to reconstruct an oversampled versions of the waveforms acquired at the radio receiver and used for ToF based localization. This method is suitable for low complexity devices, since it requires little computational resources, while at the same time offers improved localization precision by a factor of two or more. The second proposed approach is used to compensate the ranging error due to clock frequency offset in cooperative localization schemes like N-Way ranging. This approach significantly reduces the ranging and, therefore, localization errors and has much better performance compared to the existing solutions. The added system complexity is minor and the number of required wireless transmissions is kept at the theoretical minimum.

An approach for implementation of localization system in the 2.4/5 GHz ISM band is further proposed in this work. This system is implemented and tested on a software defined radio (SDR) platform. A ranging precision of better than one meter is demonstrated. Finally, an approach for integrating localization functionality into an arbitrary wireless data transmission

system is proposed. This approach is implemented in a 60 GHz wireless system which originally was designed for high data rate transmission only. A ranging precision of one centimeter with a concurrent high data rate transmission is demonstrated. No noticeable loss of ranging accuracy or data transmission performance was observed by the simultaneous operation of the two subsystems. A major benefit of the proposed approach is that it separates both functionalities, giving an efficient and flexible path to enhance existing radio systems with localization capability.

Zusammenfassung

In den letzten Jahrzehnten sind Satellitennavigationssysteme zu einem unverzichtbaren Teil des modernen Lebens geworden. Viele innovative Anwendungen bieten ortsabhängige Dienste an, welche auf diesen Navigationssystemen aufbauen. Allerdings sind diese Dienste in Innenräumen nicht verfügbar und werden durch eine urbane Umgebung negativ beeinflusst. Daher werden seit einigen Jahren alternative Lokalisierungsmethoden für Innenräume aktiv erforscht und entwickelt, um die Dienste auf die angesprochenen Bereiche zu erweitern. Hierbei oft eingesetzte Methoden sind „time of flight“ (ToF), „angle of arrival” (AoA), sowie eine Auswertung der empfangenen Signalstärke. Allgemein gesprochen sollte die Positionsgenauigkeit für Innenanwendungen deutlich höher sein als bei globalen Systemen. Andererseits sollten Komplexität und Kosten signifikant niedriger sein, da nur ein kleinerer Bereich abzudecken ist.

Der Schwerpunkt dieser Arbeit liegt darauf, die Genauigkeit von Lokalisationsmethoden in Innenräumen zu erhöhen, sowie auf der effektiven Integration der entsprechenden Verfahren in drahtlose Kommunikationssysteme. Es werden zwei Ansätze vorgeschlagen und untersucht, welche die Präzision von ToF-basierten Methoden erhöhen. Zum einen wird im „Modified Equivalent Time Sampling“ (METS) Verfahren eine überabgetastete Version der vom Radioempfänger gelieferten Wellenform erzeugt und zur ToF Bestimmung verwendet. Diese Methode eignet sich für Endgeräte mit niedriger Komplexität, da eine Erhöhung der Genauigkeit um einen Faktor zwei und mehr bereits mit vergleichsweise niedrigem Rechenaufwand erreicht wird. Der zweite erforschte Ansatz hat zum Ziel, Fehler auf Grund von Taktfrequenz-Abweichungen zu kompensieren. Dieses ist für kooperative Lokalisationsmethoden (N-Way ranging) von Bedeutung. Das in der Arbeit entwickelte Verfahren führt zu einer erheblichen Reduzierung der Fehler in der Abstandsmessung und damit der Positionsbestimmung. Dadurch liegt die Leistungsfähigkeit deutlich über der von bisher eingesetzten Lösungen. Vorteilhaft ist, dass der zusätzliche Aufwand niedrig ist und dass nur eine minimale Anzahl von Datenübertragungen benötigt wird.

Darüber hinaus wurde eine neue Methode untersucht, um Lokalisationsverfahren in Funksysteme für die ISM Bänder bei 2,4 GHz und 5 GHz zu integrieren. Die Methode wurde auf einer Software Defined Radio (SDR) Plattform implementiert und bewertet. Es konnte eine

Genauigkeit bis zu einem Meter in der Positionsbestimmung demonstriert werden. Schließlich wurde ein Verfahren vorgeschlagen und untersucht, mit welchem Lokalisationsfähigkeit in bestehende Funksysteme integriert werden kann. Die betrachtete Methode wurde in einem 60 GHz Funksystem mit hoher Datenrate implementiert. Die Untersuchungen zeigten eine Positionsgenauigkeit von 1 cm bei einer gleichzeitig hohen Datenrate für die Übertragung von Nutzdaten. Bei dem parallelen Betrieb der beiden Teilsysteme wurden weder Einbußen in der Datenübertragung noch in der Lokalisierungsgenauigkeit beobachtet. Da die Methode beide Funktionalitäten trennt, ist eine gemeinsame Integration in ein Funksystem effizient durchzuführen.

Table of contents

List of figures

List of tables

Nomenclature

Acronyms / Abbreviations

A/D	Analog to Digital
$FPGA$	Field Programmable Gate Array
SDR	Software Defined Radio
BB	Baseband
$BPSK$	Binary Phase Shift Keying
$BRAM$	Block Random Access Memory
$CDMA$	Code Division Multiple Access
CFO	Carrier Frequency Offset
$CRLB$	Cramér-Rao Lower Bound
CSS	Chirp Spread Spectrum
CW	Continuous Wave
D/A	Digital to Analog
DC	Direct Current
DDC	Digital Down-Converter
$DFCW$	Dual Frequency Continuous Wave
DFL	Device Free Localization
DP	Direct Path
$DP-NLOS$	Direct Path - Non-Line of Sight
DSO	Digital Storage Oscilloscope
$DSSS$	Direct Sequence Spread Spectrum
DUC	Digital Up-Converter
ETS	Equivalent Time Sampling
FH	Frequency Hopping
FIR	Finite Impulse Response
$FMCW$	Frequency Modulated Continuous Wave
$FPGA$	Field Programmable Gate Arrays

FSM	Finite State Machine
$GLONASS$	Global'naya Navigatsionnaya Sputnikovaya Sistema
$GNSS$	Global Navigation Satellite System
GPS	Global Positioning System
GUI	Graphical User Interface
$IR-UWB$	Impulse Radio - Ultra Wide Band
$IS-95$	Interim Standard 95
ISI	Inter-Symbol Interference
ISM	Industrial, Scientific and Medical
LCM	Least Common Multiple
LFM	Linear Frequency Modulated
LNA	Low Noise Amplifier
$LORAN$	Long Range Navigation
LOS	Line of Sight
LPS	Local Positioning System
LTE	Long Term Evolution
LTI	Linear Time Invariant
MAC	Medium Access Control
$Mcps$	Mega chips per second
MF	Matched Filter
ML	Maximum Likelihood
MLE	Maximum Likelihood Estimator
MLS	Maximum Length Sequence
MMW	Millimeter Wave
$MSps$	Mega Samples per Second
$MVUE$	Minimum-Variance Unbiased Estimator
$NDFL$	Non Device Free Localization
$NDP-NLOS$	No Direct Path - Non-Line of Sight
$NLOS$	Non-Line of Sight
$OCXO$	Oven Compensated Crystal Oscillator
$OFDM$	Orthogonal Frequency Division Multiplexing
$PAPR$	Peak to Average Power Ratio
PCR	Pulse Compression Ratio
PN	Pseudo Noise
ppm	percent per million

QAM	Quadrature Amplitude Modulation
RC	Raised Cosine
RF	Radio Frequency
RMS	Root Mean Square
RMSE	Root Mean Square Error
RSS	Received Signal Strength
RSSI	Received Signal Strength Indicator
RTT	Round Trip Time
RTToF	Round Trip Time of Flight
SDR	Software Defined Radio
$SDS-TWR$	Symmetrical Double Sided - Two Way Ranging
SF	Spreading Factor
SFDR	Spurious-Free Dynamic Range
S/N	Signal to Noise
SNR	Signal to Noise Ratio
SONAR	Sound Navigation And Ranging
SQNR	Signal to Quantization Noise Ratio
SRRC	Square Root Raised Cosine
SS	Spread Spectrum
STS	Short Training Symbols
SV	Saleh-Valenzuela wireless channel model
TCXO	Temperature Compensated Crystal Oscillator
TDD	Time Division Duplex
TDM	Time Division Multiplexing
TDoA	Time Difference of Arrival
ToA	Time of Arrival
ToF	Time of Flight
TSV	Triple S-V wireless channel model
TWR	Two Way Ranging
TWTT	Two Way Time Transfer
USRP	Universal Software Radio Peripheral
UWB	Ultra Wide Band
WiFi	Wireless Fidelity
WLAN	Wireless Local Area Network
WSN	Wireless Sensor Node

XO	Crystal Oscillator
ZZB	Ziv-Zakai Bound

Chapter 1

Introduction

1.1 Background

Location estimation of a person or an object is a centuries old idea. The main purpose of localization in the early days was its use for navigation. Polynesian navigation is one of the oldest navigation systems known. Navigators were using their own senses to estimate their current position. The knowledge about star positions, wildlife, speed and size of ocean waves was passed from one generation to the other, usually in the form of songs.

The first tool exploiting magnetic fields for navigation was the compass. It was first invented by the Chinese Han Dynasty [74], about 206 BC. The compass is still used today, especially in modern navigation systems.

Navigation systems were constantly improved. Mainly for use on seas and ships which were exploring the world or trading with remote countries. The sextant[1] was first implemented around 1730 by John Hadley and Thomas Godfrey. It was used for centuries, till new, more modern, electronic navigation and localization systems were developed.

The invention of the vacuum tube in 1904 by John Ambrose Fleming, and the development of the transistor in 1947 by John Bardeen, Walter Brattain and William Shockley, started a new era in the development of localization systems.

Nowadays, navigation and localization systems are quite modern and mainly electronic. They can be classified according to different criteria. In Figure 1.1, the basic classification of modern localization systems is presented [71]. In this thesis the main focus is on the systems using RF technology.

[1] An instrument with mirrors, used for measurement of angles between visible objects.

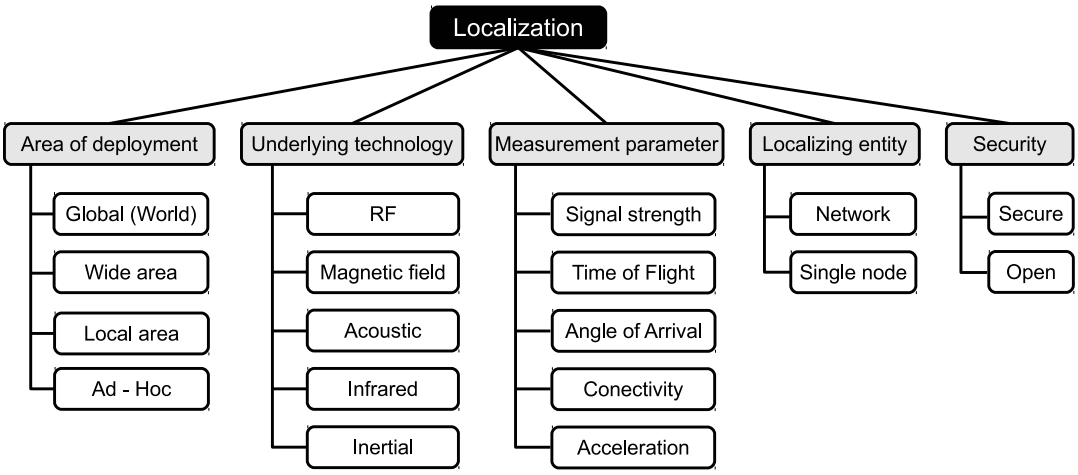


Fig. 1.1 Classification of the localization methods

One of the first RF navigation systems was the Decca Navigator System [10]. It used the hyperbolic navigation method, which is based on the comparison of the phases of two arriving signals. Another system that used the hyperbolic navigation was LORAN-C. It was introduced in 1957 and was used mainly for military purposes. Its main limitation was the expensive equipment for receiving and interpreting the signals. When the cost of the LORAN-C receiver was reduced, due to the development of integrated circuits, it became common equipment in many civil applications.

The first global-range navigation system was OMEGA. It was operated by the USA and six other nations. It became redundant with the appearance of the GPS and was later shutdown.

The GPS itself, was fully operational in 1995. It is developed, maintained and controlled by the US Department of Defense. The GPS system transmits two different signals which can be used for positioning: civilian and military. The military signal is encrypted, thus not available for civilians. On the other hand, the civilian signal is globally available and not encrypted. It is actually, the widely used global navigation system.

Meanwhile, a few other Global Navigation Satellite Systems (GNSSs) were developed and deployed. GLONASS is the Russian equivalent of the GPS. The EU developed its own system called Galileo. It is still not operational and expected to be in full service by 2020.

Two main objectives are driving the development of positioning systems. The first one is the precision. In Figure 1.2, the precision of different ranging, positioning and navigation systems is shown. The positioning precision is being constantly improved in the upcoming modern localization systems. The current GPS has a precision of less than 10 m. The second objective is achieving a larger, possibly global, coverage. GPS has global coverage, with “islands” without coverage or with reduced precision, which are normally indoor areas, urban canyons and some

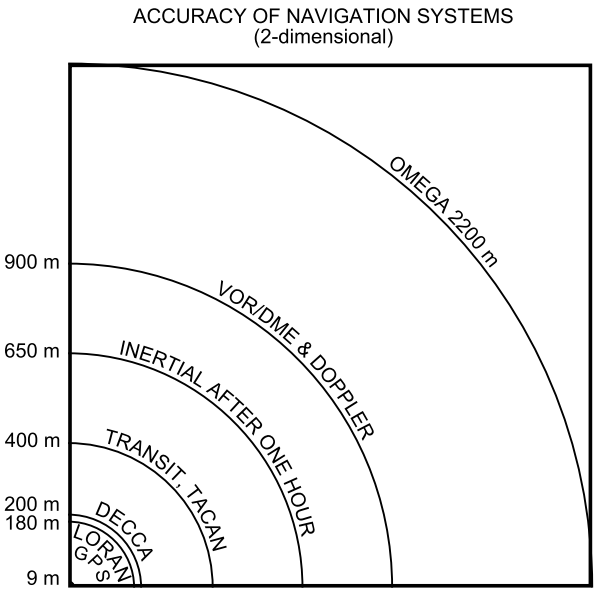


Fig. 1.2 Accuracy of navigation systems (source: Wikipedia)

special scenarios. Today, the usually accepted approach is to provide coverage to these "islands" by using other localization systems.

1.2 Motivation

The task of developing a global high-precision positioning system has been shown to be challenging. When coverage of a large area is required, usually a high precision cannot be achieved. On the other hand, systems that have higher precision, usually do not have wide area coverage. For example, the GPS system has almost global coverage, but the precision is around 10 m, or slightly less. A higher GPS precision can be obtained [49], but only in stationary scenarios or by using GPS augmentation. In a stationary scenario, the GPS receiver analyzes the incoming signals in a longer period of time, e.g. hours or even days. The precision in this case can be up to centimeter level. The other approach, GPS augmentation, uses external information in order to improve navigation attributes, such as precision, duration of measurement, availability, etc.

In the last few decades high precision low coverage systems are being actively investigated, in order to provide localization services to those areas not covered by GPS, or other GNSSs. These systems tend to be also well suited for low power and low complexity applications, which is not the case with today's GNSSs.

A common approach, for areas without GPS coverage, is to use the WiFi signal strength for localization purposes. This is possible since the power of the received signal from WiFi access points is decaying with distance, as shown in Figure 1.3. This approach is becoming interesting due to the large deployment of WiFi networks, especially indoors. The most commonly used method for WiFi based localization is called "WiFi fingerprinting". In order to use this method, a mapping of the WiFi signal strength must be performed in advance. This includes measurement of the signal strength (power level) of the surrounding WiFi access points. The measured power levels are stored in a database, together with the position where they were measured. If high precision localization is required, the signal strength measurements should be performed on many locations, densely distributed. This is usually a painstaking process, and requires significant effort. A user can obtain its location, by measuring the WiFi signal strength and compare it with the previously obtained in the mapping process. This approach can offer good localization precision [8], but if the surrounding environment is changed [2], the mapping process must be performed again.

The main problem with the "signal strength" based methods is that the received power decays fast for small distances between the wireless nodes, and slowly for large distances. Also, due to multipath propagation, the received power can exhibit fading, and would not exactly comply with the theoretical expectations, resulting from the Friis transmission equation [33]. In Figure 1.3, the theoretical and the real dependency between received power and the distance between the nodes is given. From the real power measurements, it can be seen that, due to

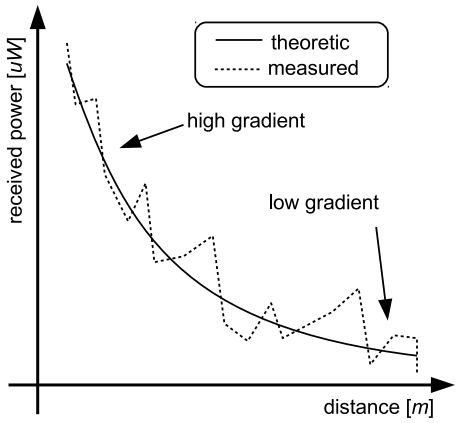


Fig. 1.3 Received signal power as a function of distance between transmitter and receiver

the fading (resulting from multipath propagation), the received power fluctuates around the theoretical value. There are no efficient methods that can be used to precisely estimate the

[2] In case there are more people in the area, or the furniture is rearranged.

distance as a function of received power in multipath propagation environments. Therefore, these methods cannot be used in localization systems where centimeter level precision is required.

High ranging and localization precision can be obtained using Time of Flight (ToF) based methods. These methods exploit the finite propagation speed of the RF signal and are more suitable for use in multipath propagation environments [36]. The main drawback is the complexity of the hardware required, as well as the associated computational complexity. Obtaining high precision distance estimates with these methods in systems with limited resources is challenging.

The main motivation for this work is to enable high precision localization of the ToF based methods when used in low complexity systems. This can be achieved by developing methods which can reduce the distance estimation errors in the ToF approach. These methods should not significantly increase the system complexity. Implementation of these methods on systems which already have wireless data transmission capability can further reduce the system cost by using the already available hardware.

The reduction of the complexity of the localization methods would enable their integration in the current and future wireless data transmission systems, leading to their wide deployment and coverage of GPS denied areas.

1.3 Scope of this work

ToF methods achieve high accuracy and can be deployed for indoor localization. However, they have high requirements in A/D and D/A converter speed, processing power and system complexity. This limits their use only in applications which are able to satisfy these requirements. This work addresses techniques that can be adopted to extend the applicability of the ToF based methods. Implementation of these techniques is also addressed, in order to verify their performance.

In Chapter 2 the current localization methods are discussed. The main focus is on trilateration, since it is a widely used positioning method. It requires distance (range) measurements in order to estimate unknown position.

The radio frequency (RF) ranging methods and waveforms are reviewed in Chapter 3. The common error sources, leading to ranging errors are also considered in this chapter. The ranging errors due to finite sample rate of the data converters, as well as errors due to frequency offset of the clock sources are presented. Approaches for reducing these ranging errors are proposed in the next chapters.

In Chapter 4 a method for reduction of the ranging error due to finite sample rate of the A/D converter is proposed. A common solution is to use complicated and costly high sample rate A/D converters in order to oversample the incoming waveform many times over the Nyquist rate. Instead, the proposed method uses A/D converters with a sample rate slightly higher than the Nyquist rate to reduce the ranging error due to finite sample rate.

Chapter 5 addresses the ranging errors due to frequency inaccuracy of the clock source. The clocks used in the WSN nodes, or other consumer devices are usually derived using crystal oscillators. The frequency of these clock sources has tolerances in the order of $\pm 10 - \pm 50$ ppm. This frequency offset is relatively small, but it can introduce large ranging errors, when methods like Two Way Ranging (TWR) are used. In cooperative localization[3] methods [85] this problem is even more pronounced. In this chapter a method for reduction of the ranging error due to clock inaccuracy in cooperative ranging scenarios is proposed and evaluated.

An approach for development of a standalone ranging system as well as integration of ranging functionality in a wireless data transmission system is described in Chapter 6. Both systems are implemented on real hardware and tested. The standalone ranging system is implemented on a Software Defined Radio (SDR) platform. The system uses the 2.4/5 GHz ISM band. This system is further extended to perform 2D localization, by estimation of distances to multiple anchor nodes. Trilateration is used to estimate the position from the distance estimates. Integration of ranging functionality in a wireless data transmission system is performed on a system utilizing the 60 GHz band. The main issues addressed here are the interfacing of the ranging modem with the MAC protocol of the data modem, as well as multiplexing the ranging frames with the data frames. Both systems, 2.4/5 GHz and 60 GHz, were tested and evaluated and the results are presented in this chapter.

The conclusion as well as the future work are given in Chapter 7.

[3] Nodes cooperate in order to estimate their relative positions

Chapter 2

Localization methods and algorithms

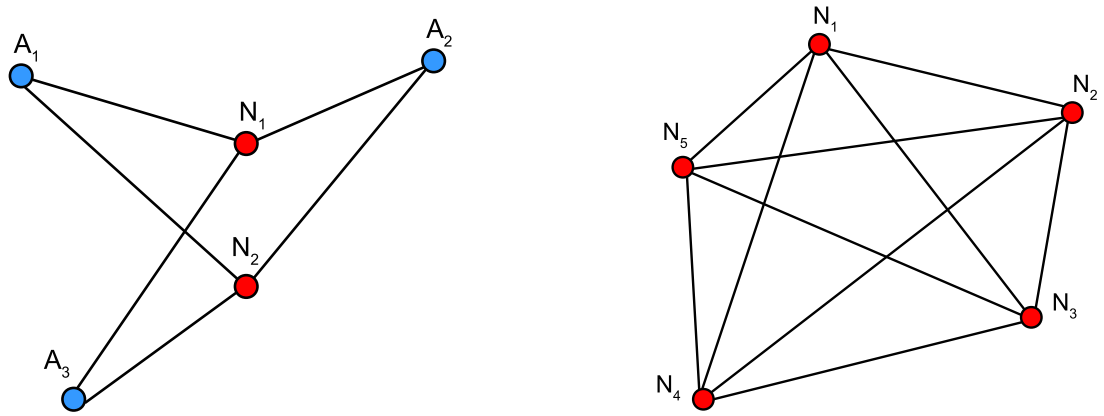
2.1 Localization - classification and methods

Localization is estimation of a location (position) of a given object in a given coordinate system. There are different means and methods to perform localization as shown in Figure 1.1. Anyway, if localization is performed using RF signals, it is commonly referred to as RF localization[1].

The RF localization methods can be classified according to different criteria. Depending on the devices used, the localization methods can be classified as **device free localization (DFL)** and **non device-free localization (NDFL)**. The **DFL** methods are used to estimate the position of an object (or human body) which does not have RF transmission capability. On the other hand, the **NDFL** methods are used to estimate the position of a device (or object) which is capable of RF transmission. In this work only the NDFL is of interest.

Depending on the deployed infrastructure the localization methods are classified as **infrastructure-based localization** and **infrastructure-free**, i.e. **cooperative localization**. In **infrastructure-based localization**, wireless devices with known position are used as a reference nodes or commonly referred to as anchors or anchor nodes. The position of the other nodes is estimated with respect to these known positions. In **Cooperative localization** no infrastructure is present. The position of a node is estimated with respect to the positions of the other nodes in a wireless network. This is a typical scenario for wireless sensor networks. In Figure 2.1, infrastructure-based and cooperative localization are presented. In the infrastructure-based localization, the anchor nodes are N_1 and N_2 and the position of the nodes A_1 and A_2 should be estimated. On the other hand, in the cooperative localization scenario no anchor nodes are present. The nodes N_1, N_2, N_3, N_4 and N_5 have unknown positions which

[1] Or RF positioning accordingly.



(a) Infrastructure based localization **(b)** Cooperative localization

Fig. 2.1 Infrastructure-based and cooperative localization

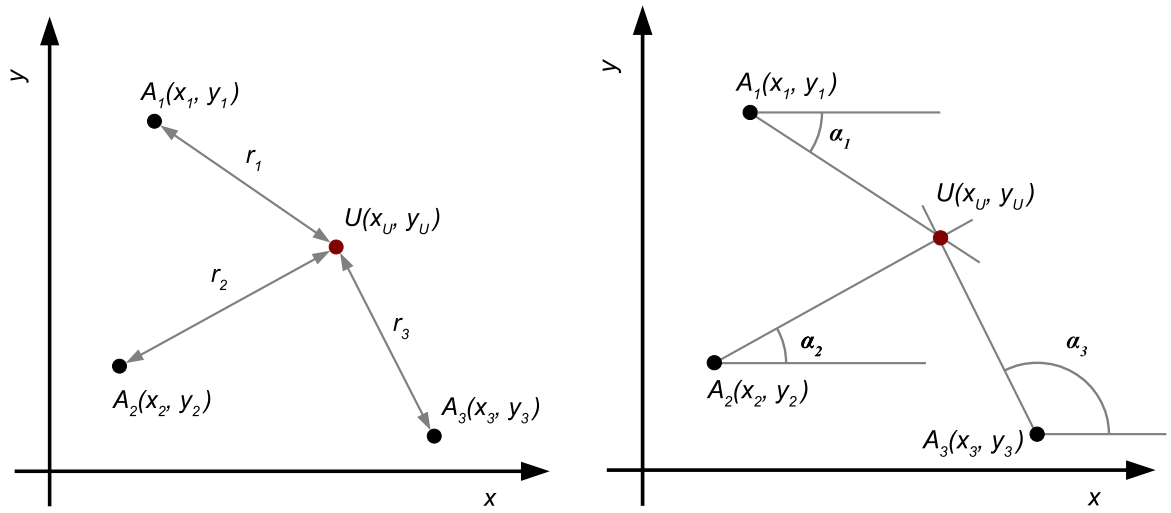
should be estimated. These position are relative to other nodes. Nevertheless, if some of the nodes have a known position, this approach can be used for absolute position estimation.

For estimation of unknown position of a node, three geometric methods are mainly used: **trilateration**, **triangulation** and **multilateration**. The commonly preferred localization method is **trilateration**, mainly due to its simplicity. The basic trilateration method is shown in Figure 2.2a. It is based on range measurements, from the unknown location to the anchor points with known coordinates. The distances (range), can be estimated by measuring different RF parameters, as received power or radio signal propagation time, i.e. time of flight.

Trilateration is also commonly used for position estimation in cooperative localization scenarios. The positions of the nodes are estimated using the distances between the nodes.

The **triangulation** method for location estimation is quite similar to trilateration. A simple illustration of the triangulation method is shown in Figure 2.2b. The unknown location is estimated by measuring the angles between the anchor points and the point with unknown location. The main complexity of triangulation comes from the need for antenna arrays in the wireless nodes to estimate the required angles. Also, most of the radios with multiple antennas (or antenna arrays), usually do not have sufficient angle resolution, needed for precise localization. Therefore, the triangulation method is usually avoided, as being too complex or imprecise, at the moment.

Multilateration is usually used in systems where only an information for the difference of the distances from the unknown location to the anchor nodes can be measured. The node residing on an unknown location measures the time difference between the arrivals of the signals from the anchor nodes. This type of localization is traditionally referred to as “hyperbolic navigation”. The name comes from the fact that when the possible locations are plotted, they



(a) Trilateration method **(b)** Triangulation method

Fig. 2.2 Trilateration and triangulation methods for location estimation

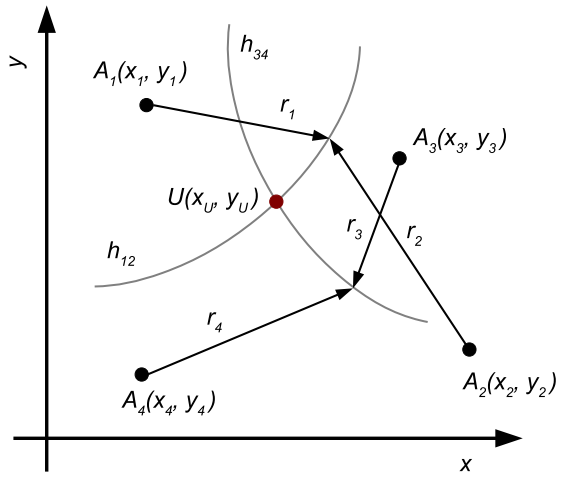


Fig. 2.3 Simplified representation of the multilateration method (two dimensional space)

form an hyperbolic curve. A simplified representation of an hyperbolic navigation system is shown on Figure 2.3. This approach is basically known as Time Difference of Arrival (TDoA) system. The most popular hyperbolic system is LORAN-C [105]. It transmits on a carrier frequency of 100 kHz, with power in order of megawatts. Due to the low carrier frequency and high transmit power, jamming[2] of this system is almost impossible. Nevertheless, although the LORAN-C should be shutdown, it is planned that some of the transmitters are still kept as a GNSS backup [41].

[2]Deliberately block or interfere with authorized wireless signal in order to disrupt the reception.

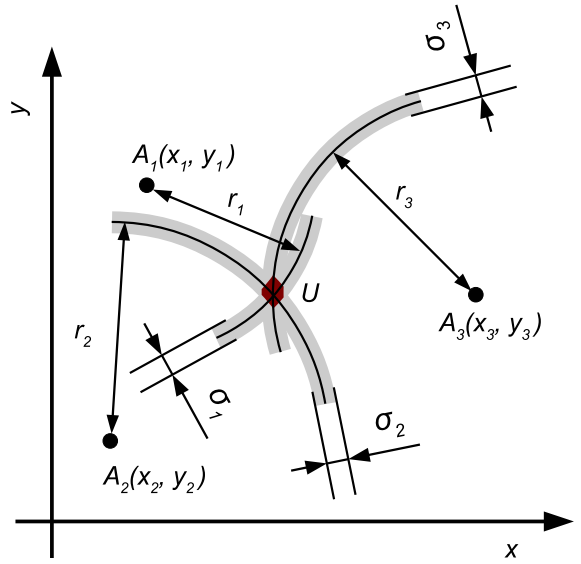


Fig. 2.4 Detailed representation of the trilateration method. Distance estimation error, σ_i, is present.

2.2 The trilateration localization method

The trilateration is the most popular method used for localization. A detailed illustration of the method is given in Figure 2.4. In order to find the unknown position of a node or an object U, the distances r_i (for $i = 1, 2, 3...$) are measured. These are the distances between the anchor nodes A_i and the node U. The range estimates include an error, marked as σ_i in Figure 2.4. Therefore, instead of estimating the exact position of node U, an area[3] containing the node is obtained. This area is marked as U in Figure 2.4.

The methods for estimation of distances between the nodes are discussed in the next chapter.

2.2.1 Position estimation using trilateration

The position of a node is estimated in a given coordinate system. For GNSSs a geodetic system[4] (or geodetic datum) is used. For indoor localization, using a Cartesian coordinate system is usually satisfactory. Also, for smaller outdoor areas, the Cartesian coordinate system is a common choice. For scenarios, where localization is to be performed in a single plane, e.g. one floor, a two dimensional Cartesian coordinate system can be used.

[3] In a three dimensional case it would be a volume instead of an area.
[4] For GPS, a most common coordinate system is WGS 84 datum (geodetic coordinate system).

Several different methods for estimating the (X, Y, Z) coordinates of a node from the r_i distance measurements are available. Some good descriptions of these methods are given in [28], [59] and [81].

In this section a derivation for a two dimensional case is shown. The main guidelines are adopted from [81], where a complete analysis for a more general three dimensional case is presented.

The position of the point U (x,y) falls in the intersection of the circles with radii r_i and centers in the (x_i,y_i), where (x_i,y_i) are the positions of the anchor nodes A_i. The equations of these circles can be written as:

$$(x-x_i)^2+(y-y_i)^2={r_i}^2 \quad (i=1,2,3,...,n) \tag{2.1}$$

These equations form a system of equations as

$$\begin{cases} (x-x_1)^2+(y-y_1)^2={r_1}^2 \\ (x-x_2)^2+(y-y_2)^2={r_2}^2 \\ \vdots \\ (x-x_n)^2+(y-y_n)^2={r_n}^2 \end{cases} \tag{2.2}$$

The solution of the system would be in the intersection points of the circles, i.e. the position of the node U. Equation (2.1), as well as the system of equations (2.2) are not linear. Therefore, in this form it is not possible to apply linear methods for solving systems of equations. In order to linearize the system of equations (2.2), x_j and y_j is added and subtracted, accordingly as

$$(x-x_j+x_j-x_i)^2+(y-y_j+y_j-y_i)^2={r_i}^2 \quad (i=1,2,3,...,n) \tag{2.3}$$

Expanding and rearranging (2.3), would lead to

$$\begin{aligned} &(x-x_i)(x_i-x_j)+(y-y_i)(y_i-y_j)= \\ =&\frac{1}{2}\left[(x-x_j)^2+(y-y_j)^2-{r_i}^2-(x_i-x_j)^2-(y_i-y_j)^2\right]= \\ =&\frac{1}{2}\left[{r_j}^2-{r_i}^2+{r_{ij}}^2\right]=b_{ij} \end{aligned} \tag{2.4}$$

In (2.4), r_i is the distance form node U to anchor node A_i, r_j is the distance from U to A_j, and r_{ij} is the distance between the anchor nodes A_i and A_j. The distance r_{ij} can be calculates as

$$r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \tag{2.5}$$

Equation (2.4) can be further written as

$$(x - x_i)(x_i - x_j) + (y - y_i)(y_i - y_j) = \frac{1}{2}\left[{r_j}^2 - {r_i}^2 + {r_{ij}}^2\right] = b_{ij} \tag{2.6}$$

Using (2.6) for all anchor nodes, and arbitrary fixing j and $j \neq i$, the following system of equations would be obtained

$$\begin{cases} (x - x_2)(x_2 - x_1) + (y - y_2)(y_2 - y_1) = \frac{1}{2}\left[{r_1}^2 - {r_2}^2 + {r_{21}}^2\right] = b_{21} \\ (x - x_3)(x_3 - x_1) + (y - y_3)(y_3 - y_1) = \frac{1}{2}\left[{r_1}^2 - {r_3}^2 + {r_{31}}^2\right] = b_{31j} \\ \vdots \\ (x - x_n)(x_n - x_1) + (y - y_n)(y_n - y_1) = \frac{1}{2}\left[{r_1}^2 - {r_n}^2 + {r_{n1}}^2\right] = b_{n1} \end{cases} \tag{2.7}$$

In the case in Figure 2.4, only the first two equations would exist. The system can be written in matrix form as

$$Ax = b \tag{2.8}$$

where

$$A = \begin{pmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{pmatrix} \quad x = \begin{pmatrix} x - x_1 \\ x - x_2 \end{pmatrix} \quad b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \tag{2.9}$$

It can be noticed that there are two unknowns, the x and y coordinates of node U. There are also two equations, which result in a fully determined system, if the two equations are linearly independent.

The equation (2.8) can be solved by simply finding A^{-1}, which is computationally intensive. Usually Gaussian elimination and LU factorization if A is square or QR factorization [107] if A is rectangular are used, since they require less computational effort.

2.2.2 Trilateration with non error free distance measurements

The estimates of the distances r_i contain an error, due to noise and different imperfections of the ranging system. Therefore, the left and the right side of the equation (2.8) are not anymore

equal. A more appropriate form of equation (2.8), in this case, would be

$$Ax \approx b \tag{2.10}$$

According to (2.10) the circles with radii r_i around the anchor nodes would not intersect in a single point.

A standard approach for approximately solving linear systems of equations, like the one in (2.10), is the linear least squares method [9]. This method basically minimizes the sum of the residuals r of the given linear system.

$$r(x) = (b - Ax) \tag{2.11}$$

The sum of the residuals would be

$$S(x) = \sum_{i=1}^{m} |b_i - \sum_{j=1}^{n} A_{ij}x_j|^2 = \|b - Ax\|^2 \tag{2.12}$$

and the actual problem of estimating the coordinates can be written as

$$\hat{x} = \underset{x}{argmin}\ S(x) \tag{2.13}$$

where $\hat{x}$ are the estimated coordinates.

The sum of residuals can be further expanded as

$$S(x) = \|b - Ax\|^2 = (b - Ax)^T(b - Ax) = b^T b - x^t A^T b - b^T Ax + x^T A^T Ax \tag{2.14}$$

using the matrix form.

In order to minimize the sum $S(x)$, the first derivative of equation (2.14) should be found and set to 0.

$$\frac{\partial S(x)}{\partial x} = 2A^T Ax - 2A^T b = 0 \tag{2.15}$$

Equation (2.15) is called normal equation and is usually written as

$$A^T Ax = A^T b \tag{2.16}$$

If $A^T A$ is substituted with A_N, and $A^T b$ with b_N, again a system of equations $A_N x = b_N$ is to be solved. There are standard methods for solving this system of equations, but they are not going to be discussed here since they are beyond the scope of this work.

Chapter 3

Radio frequency ranging methods and limitations

This chapter is devoted to RF ranging methods used for estimation of the distance between the nodes. The main focus is on the Time of Flight (ToF) based methods, but also Received Signal Strength (RSS) based ranging methods are discussed. Besides, waveforms used for ToF ranging are presented.

The error sources as well as the ranging precision limitations are also discussed in this chapter. The main contribution of this thesis is focused on reducing and compensation of some of the errors described here.

This chapter is organized as follows: in Section 3.1 an overview of the RSS based distance estimation methods is presented; Section 3.2 is devoted to the ToF based ranging methods; in section 3.3 an overview of the commonly used waveforms is given; finally, Section 3.4 presents the common ranging error sources and distance estimation precision limitations.

3.1 Received signal strength (RSS) based distance estimation

The simplest approach for RF based distance estimation uses the received RF power level (or RSS), which decreases by increasing the distance between the transmitter and the receiver. The Friis equation (3.1) is used to estimate the received signal power level, i.e. RSS [91], as a function of distance.

$$P_r^{dB}(d) = P_0^{dB} - 10n\log\frac{d}{d_0} + X_\sigma^{dB} \quad [dB] \tag{3.1}$$

In (3.1) the coefficient n is called "path-loss exponent". It is 2 for free space and usually higher for multipath propagation environments [2], e.g. indoors. The received signal power in dB is $P_r^{dB}(d) = 10\log_{10}(P_r(d)/1\ W)$. $P_r(d)$ is the received signal power measured in watts at a distance of d meters from the transmitter. P_0^{dB} is the received power (in dB) at distance d_0. The $X_\sigma^{dB} = 10\log_{10}(X_\sigma)$ is the noise power at the receiver. X_σ has a log normal distribution as shown in [19, 45, 91]. Therefore, X_σ^{dB} expressed in dB would also have normal distribution

$$X_\sigma^{dB} \sim N(0, \sigma_r^2) \tag{3.2}$$

where σ_r is the standard deviation of the received noise power.

The received signal power $P_r^{dB}(d)$ is also normally distributed

$$P_r^{dB}(d) \sim N(\overline{P}_r^{dB}(d), \sigma_r^2) \tag{3.3}$$

where $\overline{P}_r^{dB}(d)$ is the mean received power.

Distance estimation from the RSS measurements is usually performed using a Maximum Likelihood Estimator (MLE) [84]. The likelihood, or the density function, $f(P_r^{dB}|d)$, can be expressed as

$$f(P_r^{dB}|d) = N(\overline{P}_r^{dB}(d), \sigma_r^2) = \frac{1}{\sigma_r\sqrt{2\pi}} \exp \frac{(P_r^{dB} - \overline{P}_r^{dB}(d))^2}{2\sigma_r^2} \tag{3.4}$$

The log-likelihood is defined as

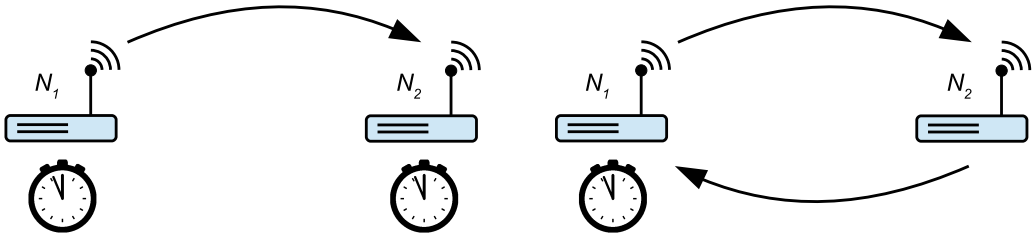
$$\ln f(P_r^{dB}|d) = -\frac{1}{2}\ln(2\pi\sigma_r^2) - \frac{1}{2\sigma_r^2}(P_r^{dB} - \overline{P}_r^{dB}(d))^2 \tag{3.5}$$

To find the maximum of the likelihood function, it should be differentiated with respect to $\overline{P}_r^{dB}(d)$ and set to zero. Therefore, the maximum likelihood is achieved for $P_r^{dB} = \overline{P}_r^{dB}(d)$. The maximum likelihood estimated distance d_{MLE}, i.e. the distance between the nodes, would be

$$d_{MLE} = d_0 10^{\frac{P_0^{dB} - P_r^{dB}(d)}{10n}} \tag{3.6}$$

which is equal to the inversion of Friis equation (3.1) The distance d_{MLE} has a log-normal distribution. According to [84]

$$E\left[d_{MLE}\right] = Cd \tag{3.7}$$



(a) Time of Arrival - Both nodes must have precisely synchronized clocks.

(b) Two Way Ranging - Only a coarse synchronization between the nodes is necessary.

Fig. 3.1 Time of Flight based ranging methods

where $C = e^{\frac{\gamma}{2}}$ and γ is

$$\gamma = \left(\frac{10n}{\sigma_r^{dB} \log 10} \right)^2 \tag{3.8}$$

Using the channel model from [91], the value of C can be estimated to be $C \approx 1.2$. This represents a 20% error and, as proposed by [84], the MLE should incorporate this in order to minimize the error. This error is also proportional to the distance.

The RSS ranging method does not require additional hardware in the wireless nodes (devices), except the already available wireless data transceiver. This transceiver usually has a RSS circuit which measures the received power level in order to perform signal conditioning (amplification). The RSS information is also available to the user and is called Received Signal Strength Indicator (RSSI). Therefore, this method can be easily implemented in wireless nodes.

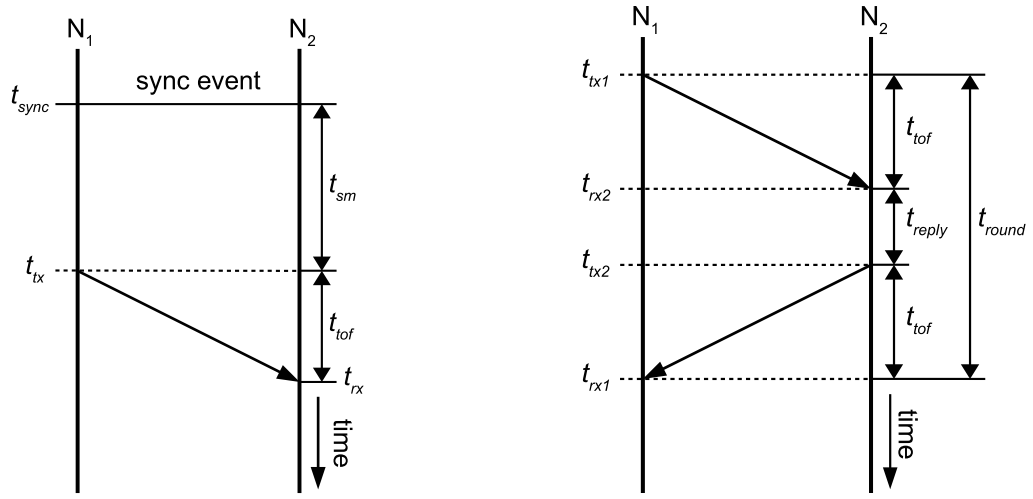
3.2 Time of flight based distance estimation

The ToF based ranging methods exploit the RF wave propagation delay between two wireless nodes to estimate the distance. The propagation speed of RF wave (in air) is equal to the speed of light c and is approximately 299792458 m/s.

A few different methods are used for ToF based distance estimation. The two most common are **Time of Arrival (ToA)** and **Two Way Ranging (TWR)**, shown in Figure 3.1.

3.2.1 The Time of Arrival (ToA) method

The ToA method shown in Figure 3.1a is usually used in GNSS systems [58]. The main advantage of this method is its precision. The main disadvantage is the requirement of high precision synchronized clocks in the nodes, which is challenging and costly. Therefore, this



(a) Timing diagram of time of arrival method

(b) Timing diagram of Two Way Ranging method

Fig. 3.2 Timing diagrams of Time of Flight based ranging methods

method is usually avoided, except in GNSS. The GPS, for example, has an atomic clock in every satellite, which is usually not suitable for implementation in small and portable devices.

A timing diagram of the ToA method is shown in Figure 3.2a. The vertical lines represent the time, one for the node N_1 and one for N_2. The arrows represent the transmissions. The *sync event* represents the performed synchronization of the clocks of both nodes. To estimate the distance between the nodes N_1 and N_2 using this method, a waveform is transmitted from N_1 to N_2, as shown. The node N_2 receives the waveform and estimates the time of arrival. The time the waveform was transmitted, t_{tx}, is known to both nodes. It was agreed by the nodes in advance. The time of flight is estimated as a difference between the time of arrival, t_{rx} and the time of transmission, t_{tx}. The distance is calculated by multiplying the time of flight by the speed of light. For example, if the signal was transmitted by N_1 at time $t_{tx} = 0$ and it was received by N_2 at time $t_{rx} = 23\ ns$, the time of flight would be $t_{tof} = t_{rx} - t_{tx} = 23\ ns$. The estimated distance would be $d = t_{tof}c = 6.9\ m$.

3.2.2 The Two Way Ranging (TWR) method

The most popular ToF based distance estimation method is the TWR. It is a common choice in indoor localization applications.

Two transmissions are needed to perform TWR. Node N_1 transmits a waveform to node N_2, which replies back to node N_1, as shown in Figure 3.1b. The time of flight from node N_1 to node N_2 is t_{tof}, as shown on the timing diagram in Figure 3.2b. The time needed for node N_2 to

reply to node N_1 is t_{reply}. The overall time for transmission of the signal from node N_1 to node N_2 and receiving it back to node N_1 is the round trip time, t_{round}, calculated as

$$t_{round} = t_{tof} + t_{round} + t_{tof} = 2t_{tof} + t_{reply} \tag{3.9}$$

By rearranging (3.9), the t_{tof} can be calculated as

$$t_{of} = \frac{1}{2}\left(t_{round} - t_{reply}\right) \tag{3.10}$$

Equation (3.10) is used for estimation of the time of flight and thus the distance between the nodes N_1 and N_2. The distance is further calculated by multiplying the time of flight by the speed of light.

The TWR methods require only coarse or no synchronization between the nodes. The coarse synchronization might be required only in systems where the nodes are supposed to allocate a special time window [25] for performing the TWR. This synchronization usually should not have better precision than a few microseconds [25] and is easily achievable. Some implementations of the TWR method do not require any synchronization between the nodes. For example, [67] uses a special burst mode which in advance takes into account that the nodes are not synchronized. The other approach is shown in [83]. In this case a synchronizer is used to detect the incoming ranging waveform. However, this is only a simple coarse synchronizer, which triggers the ranging procedures.

3.3 Waveforms used in ToF based ranging

As described previously, the problem of distance estimation between two wireless nodes is a problem of ToF estimation. To perform ToF estimation using the ToA or TWR method, the time of reception (arrival) must be estimated at the receiving nodes. For the ToA method, the time of arrival is estimated once and in the TWR method twice. The ToA method and the ToA estimation should not be confused. The ToA method refers to the method itself, described in Section 3.2.1, where ToA estimation refers to the estimation process.

This section is devoted to the waveforms used for ToA estimation at the receiver. A few different waveform types are commonly used for this purpose. Depending on the used wireless system as well as the scenario, some of the waveforms can be better suited than the others. The main advantages and disadvantages are presented here. In this section three different waveform

types used for ToA estimation are presented. They include: **short pulses**, **sine waveforms** and **pulse compression waveforms**.

3.3.1 Short pulses

The simplest waveforms used for ToA estimation are short pulses. To perform a ToA estimation, one node sends a short pulse to the other node. The other node detects the arrival of the pulse using a simple threshold detector. The threshold can be fixed to a certain value or it can be adaptively changed.

Ranging using short pulses is mainly used in impulse radio - ultra wide-band (IR-UWB) systems. Due to the large bandwidth used, it achieves centimeter precision [65, 104]. The IR technology does not require high speed data converters and complex baseband processors. Instead, simple high speed pulse generators and pulse detectors are used. The main limitation of the IR-UWB technology is the limited transmission power. Since the UWB band is overlapping with other bands, the power transmitted by UWB devices must be limited. UWB transmissions are seen as noise by the other devices operating in bands overlapping with UWB. Usually, the UWB devices are intended to work for distances of up to 10 meters [65], but there are also some devices which can reach up to a few hundreds of meters [32]. Due to the low transmit power, this system is not a good candidate for use in low SNR scenarios.

3.3.2 Sine waveforms

Sine waveforms are the simplest waveforms which can be used in ranging applications for ToA estimation. They are usually used in RADAR systems, where the distance to a passive target is estimated. The ToA in RADAR systems is estimated using the phase or the frequency of the received sine waveform, being reflected from the target.

In this section Continuous Wave (CW) and Dual Frequency Continuous Wave (DFCW) RADAR waveforms are discussed for use in indoor positioning systems [86, 113]. The use of these waveforms for ToA estimation does not require complex hardware, but it can be sensitive to multipath propagation. These waveforms are almost exclusively used only in TWR methods, where the phase of the received sine waves represents the ToA.

Continuous Wave (CW) RADAR waveforms

The use of CW RADAR waveforms in TWR method is shown in Figure 3.3. The node N_1 sends a sine wave to node N_2. Node N_2 synchronizes to this sine wave and transmits back a

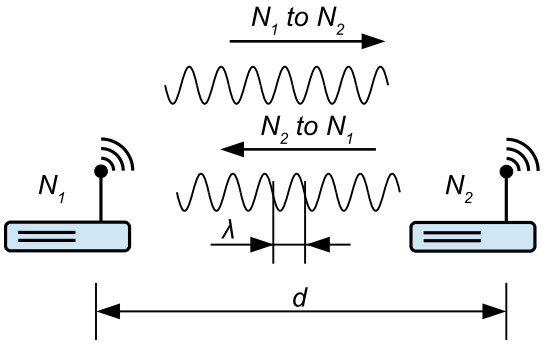


Fig. 3.3 Continuous wave RADAR method, used for two way ranging. The nodes transmit sine waves.

sine wave with the same frequency and phase. Subsequently, node N_1 compares the phases of the sent and received waveform and estimates the distance. Assuming that a sine wave with a frequency ω, amplitude A and phase φ is transmitted, the signal at the output of the transmitter would be

$$y_1(t) = Asin(\omega t + \varphi). \tag{3.11}$$

The transmitted signal travels a distance of $2d$ from node N_1 to node N_2 and back. The waveform received at node N_1 is represented as

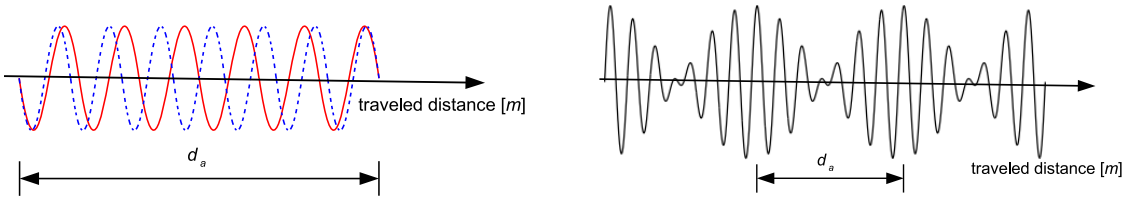
$$y_2(t) = Asin(2kd \pm \omega t + \varphi) \tag{3.12}$$

where $k = \frac{2\pi}{\lambda}$. The receiver estimates the phase of the received waveform $y_2(t)$ and compares it to the phase of the transmitted waveform, $y_1(t)$. The phase difference between them is given by $\varphi_{12} = 2kd \pm 2n\pi$. The factor $2n\pi$, where $n = 2d/\lambda$ (the symbol "/" denotes integer division), represents the phase ambiguity in the phase difference estimation. The estimated phase of the received waveform is proportional to the ToA at node N_1 and the phase difference between the transmitted and received waveforms is proportional to the round trip time. The distance d can be calculated as

$$d = \lambda \left(\frac{\varphi_{12}}{2\pi} \pm n \right). \tag{3.13}$$

In systems that utilize the 2.4 GHz and 5 GHz ISM bands, the wavelengths and, therefore, the ambiguity distances are $\lambda_{2.4G} \approx 12.5\ cm$ and $\lambda_{5G} \approx 5\ cm$ respectively. Performing a CW based TWR at these frequencies is not applicable for indoor ranging scenarios. Alternatively, this method can be used for very precise[1] tracking [5]. By continuous tracking of the phase

[1] Precision of $\approx 1\ mm$ has been demonstrated.



(a) Two sine waves with slightly different frequencies

(b) The waveform obtained by summing two sine waves

Fig. 3.4 The Dual Frequency Continuous Wave (DFCW) method waveform consists of two sine waves

difference, the change of the distance between the wireless nodes can be tracked. Nevertheless, the absolute distance in this case cannot be estimated.

Dual Frequency Continuous Wave (DFCW) RADAR waveforms

In order to mitigate the phase ambiguity problem, waveforms used in DFCW RADAR are a common choice in ranging and positioning applications. These waveforms consist of two sine waves with slightly different frequencies. In Figure 3.4, these two sine waves are shown.

Due to different frequencies of the sine waves, the phase difference between them would change with distance. It would be equal at multiples of distance d_a, which is also referred to as an ambiguity distance. The ambiguity distance is actually the least common multiple of λ_1 and λ_2, i.e. $d_a = lcm(\lambda_1, \lambda_2)$.

For different choices of frequencies, different ambiguity distances (d_a) can be achieved, as shown in Table 3.1. For the frequencies in the last row in Table 3.1, the ambiguity distance is $\approx 30\ m$, which is usually sufficient for a large number of indoor applications.

This approach can be further extended by using multiple frequencies, which main advantage is larger ambiguity distance and better ranging precision. A similar approach is presented in [113].

One common problem which affects both the CW and the DFCW based ranging systems is the multipath propagation of the transmitted sine waveforms. In indoor scenarios this effect is extremely pronounced. Mitigating, or even detecting the ranging error due to multipath propagation, can be a challenging task with this approach. For example, if a waveform $y(t) = A\sin(\omega t + \varphi)$ is transmitted, in a multipath propagation environment, more than one sine waves would arrive at the receiver. Assuming that the direct path (DP) wave is $y_1(t) = A_1\sin(kd_1 + \omega t + \varphi_1)$ and one of the reflected is $y_2(t) = A_2\sin(kd_2 + \omega t + \varphi_2)$, the waveform obtained at the receiver would be $y_R(t) = A_R\sin(kd_R + \omega t + \varphi_R)$. The distance estimate d_R

Table 3.1 Least common multiple of λ_1 and λ_2 for different frequencies in the 2.4 GHz ISM band

f_1 [GHz]	λ_1 [cm]	f_2 [GHz]	λ_2 [cm]	LCM of λ_2 and λ_2 [cm]
2.400	12.49	2.480	12.09	37.47
2.400	12.49	2.440	12.29	74.95
2.420	12.39	2.440	12.09	149.90
2.420	12.39	2.430	12.33	299.79
2.411	12.43	2.478	12.09	2997.92

would contain an error due to the reflected wave. It is not possible at this stage to know that the range estimation contains an error due to the reflection received.

To avoid the ranging error due to multipath propagation, multiple measurements on different frequencies must be performed. For each frequency, or pair (multiple) of frequencies, the DP wave and the reflected wave are combined with a different phase difference at the receiver. This approach is used in [113]. Given the large number of measurements required, the overall process usually lasts longer compared to other ranging methods.

3.3.3 Analog pulse compression waveforms

Pulse compression is a signal processing method which is mainly used in RADAR, SONAR and in ultrasound diagnostics. This method is used to increase the ranging resolution, especially in low SNR scenarios, by compressing the pulse duration. Using signal processing techniques, it produces short pulses from long waveforms. It is also a common choice in different RF ranging and positioning applications [30].

Better spatial resolution, using ToF based methods, is achieved by transmission of very short pulses. These pulses must have large amplitudes in order to achieve high SNR at the receiver. A waveform composed of short pulses with large amplitudes has a high peak-to-average power ratio (PAPR). In Figure 3.5, a simple comparison of two pulses with high and low PAPR is shown. In the narrow red pulse the power is concentrated in a short time interval, leading to high PAPR. On the other hand, the wide blue pulse has significantly smaller instantaneous power, but dissipated in a larger time interval, i.e. low PAPR. Nevertheless, both pulses contain the same energy[2]. Designing and implementing a RF output power amplifier for waveforms with high PAPR is challenging and usually avoided.

Pulse compression waveforms have nearly constant instantaneous power over time and, therefore, low PAPR. The pulse compression waveforms occupy a bandwidth of B and have a

[2]The energy is calculated as $P\Delta t$, where P is the transmitted power and the Δt is the pulse duration.

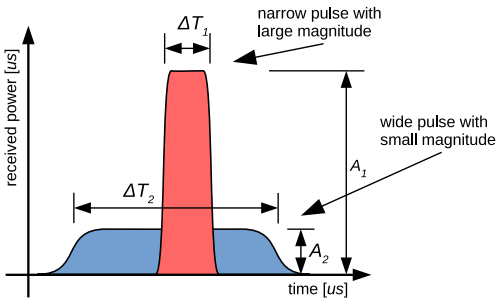


Fig. 3.5 The wide and the narrow pulse have the same energy but different duration

duration of τ. Their time-bandwidth product $\tau B \gg 1$ is large, compared to short pulses, which have $\tau B \approx 1$. The larger time-bandwidth product leads to better noise immunity as well as improved spatial resolution.

When pulse compression waveforms are used in ToF based ranging, the receiver must estimate the ToA of the incoming waveform. To obtain narrow pulses needed for precise ToA estimation and, therefore, good spatial resolution, the received waveform must be filtered with a matched filter. The received waveform consists of the transmitted waveform $s(t)$ attenuated by a factor of a and thermal noise $w(t)$. It can be represented as $r(t) = a\, s(t) + w(t)$. The matched filtering is performed as a cross-correlation of the received waveform with a locally generated copy of the transmitted waveform as

$$(s * r)(\tau) = \int_0^\infty s(t) r(t-\tau) dt \tag{3.14}$$

By substituting the expression for $r(t)$ in (3.14) the cross-correlation would be

$$\begin{aligned}(s * r)(\tau) &= \int_0^\infty s(t)(a\, s(t-\tau) + w(t-\tau)) dt \\ &= a(s * s)(\tau) + (s * w)(\tau).\end{aligned} \tag{3.15}$$

In (3.15), the obtained cross-correlation function is a sum of the autocorrelation function of $s(t)$, $(s * s)(\tau)$, and cross-correlation of $s(t)$ with the noise present at the receiver, $(s * w)(\tau)$. The cross-correlation of the noise and the waveform $s(t)$ should be theoretically equal to zero, since the noise is white and not correlated with the transmitted waveform.

The output of the matched filter, $(s * r)(\tau)$, should contain a narrow peak in order to provide good timing resolution when the time of arrival is estimated. This narrow peak can be only a part of the autocorrelation function of $s(t)$. Therefore, waveforms having a narrow peak in their autocorrelation function, for $\tau = 0$, are good candidates for use in pulse compression methods.

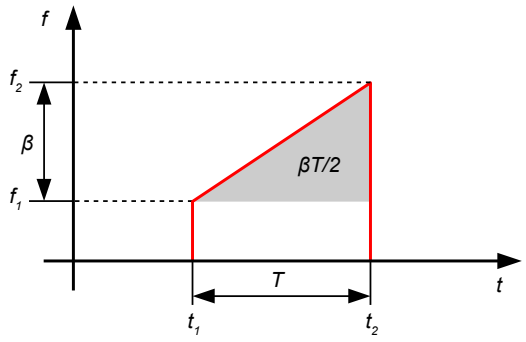


Fig. 3.6 Frequency as a function of time of a chirp pulse

Large number of waveforms have an autocorrelation function containing a narrow peak and can be used as pulse compression waveforms. The simplest analog pulse compression waveforms used in RADAR and in ranging applications [87] are the linear frequency modulated (LFM) waveforms.

Linear Frequency Modulated (LFM) waveforms

A common waveform used for pulse compression is the linear frequency modulated[3] waveform or pulse. The LFM pulse is a sine wave with with duration of T seconds, fixed amplitude and linearly varied frequency over time. Figure 3.6 depicts the frequency variation over time of a LFM pulse. The LFM pulse starts at time t_1 and ends at time t_2. The frequency changes linearly, starting from f_1 at time t_1 and ending with f_2 at time t_2. The LFM pulse envelope is given by

$$c(t) = \begin{cases} A\cos(2\pi\left[f_1 + \frac{f_2-f_1}{t_2-t_1}(t-t_1)\right]t) & for \quad t_1 \leq t \leq t_2 \\ 0 & for \quad t < t1 \quad or \quad t > t2 \end{cases} \tag{3.16}$$

According to [102], the autocorrelation function of the LFM pulse is

$$R_{LFM}(\tau) = \begin{cases} \frac{A^2}{2}\frac{\sin[\alpha\tau(T-|\tau|)/2]}{\alpha\tau(T-|\tau|)/2}(T-|\tau|)cos\omega_0\tau & for \quad -T \leq \tau \leq T \\ 0 & for \quad \tau < -T \quad or \quad \tau > T \end{cases} \tag{3.17}$$

where $\alpha = \frac{f_2-f_1}{t_2-t_1} = \frac{\beta}{T}$ is the frequency slope of the LFM pulse, A is the amplitude of the LFM pulse from (3.16), $T = t_2 - t_1$ is the length of the pulse and $\beta = f_2 - f_1$ is the band in which the frequency of the LFM pulse is varied. The frequency $\omega_0 = 2\pi\frac{f_1+f_2}{2}$ is the carrier frequency at which the LFM pulse is transmitted.

[3] Also commonly referred to as chirp signal or chirp pulse.

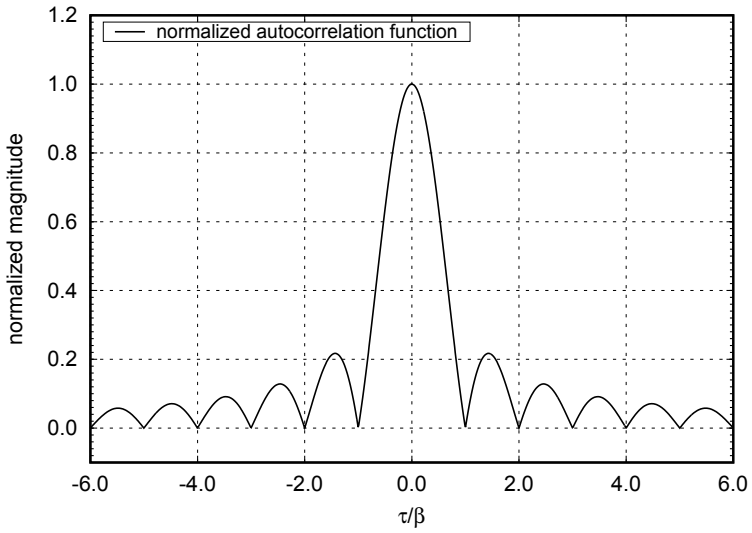


Fig. 3.7 Autocorrelation function of a LFM (chirp) pulse

The plot of the autocorrelation function of the LFM pulse is shown in Figure 3.7. It shows the magnitude of the autocorrelation function from (3.17), with τ being normalized by β. The autocorrelation function (3.17), derived in [21], is a *sinc* function on the interval $[-T,T]$ and zero for $|\tau| > T$. The width of the central autocorrelation pulse depends only on β, i.e. the difference $f_2 - f_1$, as shown in Figure 3.7.

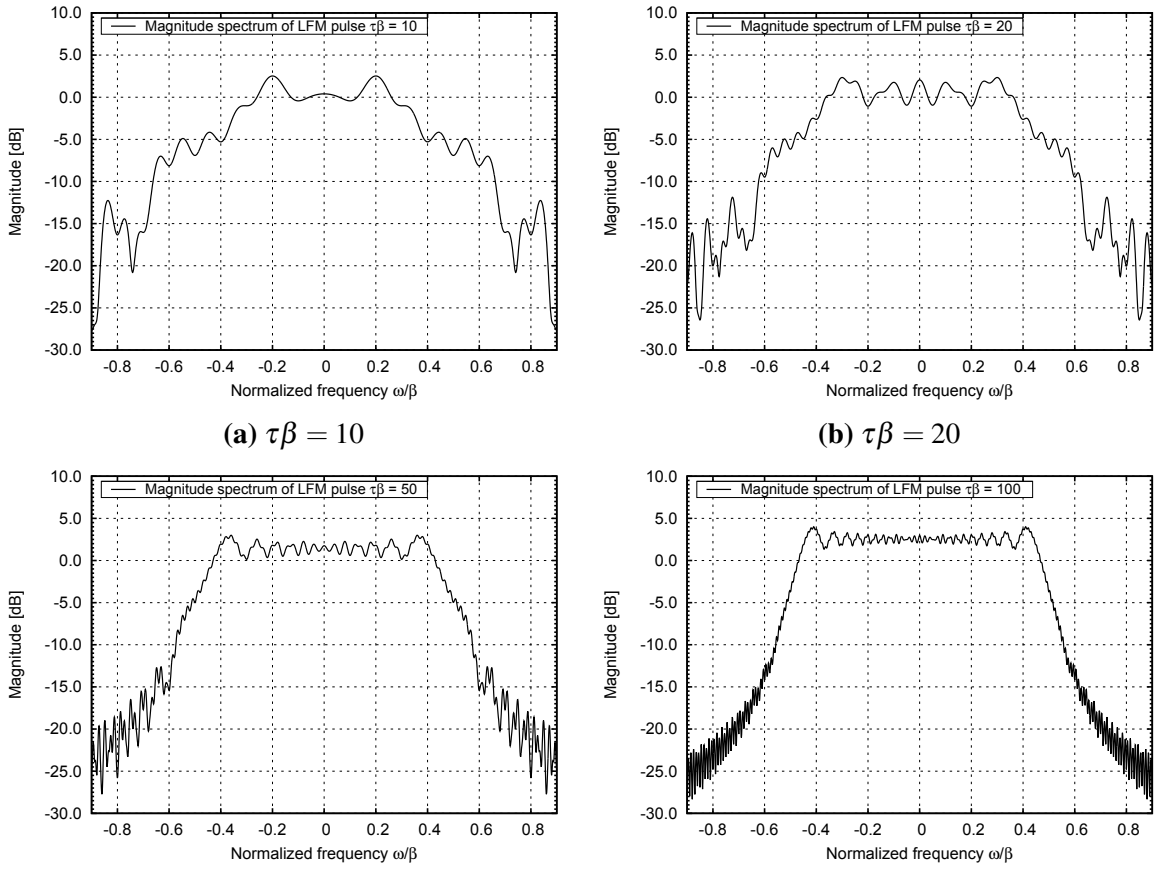
Since the frequency of the LFM pulse changes from f_1 to f_2, it is natural to assume that the occupied bandwidth B is equal to $\beta = f_2 - f_1$. This is approximately true, but precise calculation of the frequency spectrum and, therefore, the power spectral density of a LFM pulse, can give better insight of the occupied bandwidth. In [102], a complete and detailed derivation of the LFM pulse spectrum is performed and the magnitude of the spectral density is given as

$$|S(\omega)| = \frac{A}{2}\sqrt{\frac{\pi}{\alpha}}\sqrt{\left[C_f(\nu_1)+C_f(\nu_2)\right]^2+\left[S_f(\nu_1)+S_f(\nu_2)\right]^2}. \tag{3.18}$$

The functions $C_f(x)$ and $S_f(x)$ are Fresnel[4] integrals, calculated as

$$\begin{aligned} C_f(x) &= \int_0^x \cos\frac{\pi\xi^2}{2}d\xi \\ S_f(x) &= \int_0^x \sin\frac{\pi\xi^2}{2}d\xi. \end{aligned} \tag{3.19}$$

[4] Augustin Jean Fresnel, French mathematician, 10 May 1788 - 14 July 1827.



(a) $\tau\beta = 10$ **(b)** $\tau\beta = 20$

(c) $\tau\beta = 50$ **(d)** $\tau\beta = 100$

Fig. 3.8 Magnitude spectrum of a LFM pulse for different values of $\tau\beta$

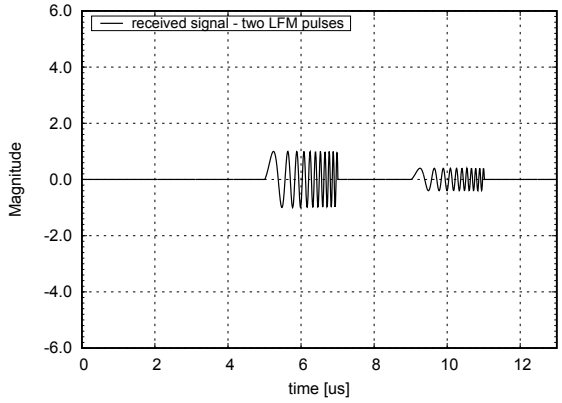
The expressions in (3.19) do not have closed form solution, but can be evaluated numerically or can be expanded using the Taylor power series. The arguments v_1 and v_2, from (3.18) are given as

$$\begin{aligned} v_1 &= \frac{\alpha\tau + 2(\omega - \omega_0)}{2\sqrt{\alpha\pi}} \\ v_2 &= \frac{\alpha\tau - 2(\omega - \omega_0)}{2\sqrt{\alpha\pi}} \end{aligned} \tag{3.20}$$

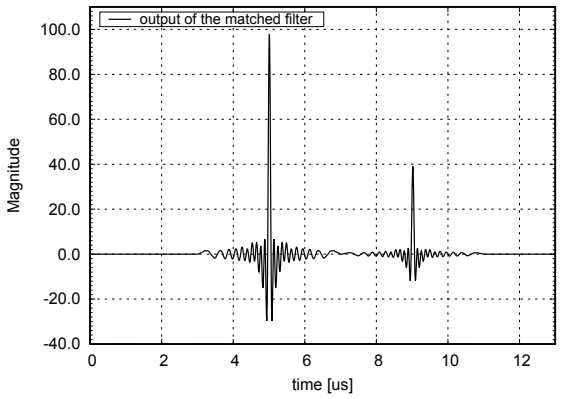
where $\omega = 2\pi f$ and $\omega_0 = 2\pi\frac{f_1+f_2}{2}$.

The magnitude part of the spectrum is shown in Figure 3.8. The bandwidth, B, of the LFM pulse is approximately equal to $\beta = f_2 - f_1$, as initially assumed.

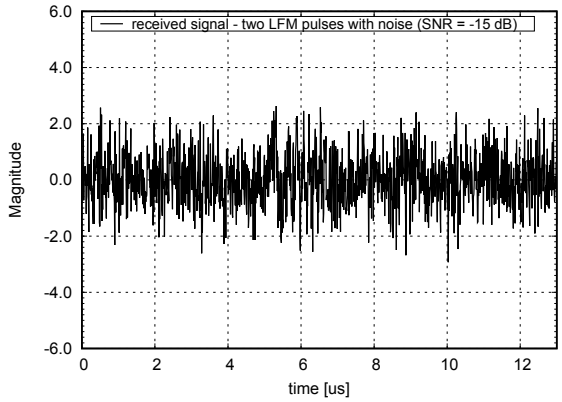
To demonstrate the pulse compression method using LFM pulses, a simulation was performed and the results are shown in Figure 3.9. It was assumed that a single LFM pulse is transmitted and that two pulses are received: one traveling the direct path and the other being reflected from an obstacle. Two cases are examined.



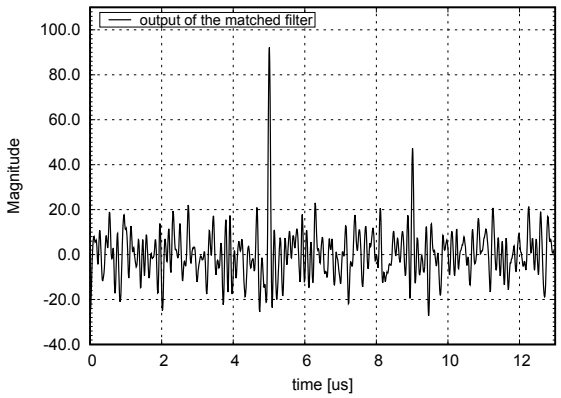
(a) Two LFM pulses received at the receiver. No added noise.



(b) Output of the matched filter for the signal in a)



(c) Two LFM pulses with additional noise ($SNR = -15\ dB$)



(d) Output of the matched filter for the signal in c)

Fig. 3.9 Pulse compression using LFM pulses

In the first case, Figure 3.9a, two LFM pulses are received and no additional noise is present. Figure 3.9b depicts the waveform at the output of the matched filter. It consists of two narrow pulses, with peaks being at the same position as the start of each LFM pulse.

The second case is more realistic, since it has a significant amount of noise included. In this case, the SNR is $-15\ dB$. In the waveform shown in Figure 3.9c, the LFM pulses cannot be observed, due to the large amount of noise. In Figure 3.9d, the output of the matched filter is shown. The two expected peaks can be clearly distinguished. Having a significant amount of noise in the signal does not significantly affect the detection of the pulses after the matched filter. This property of the pulse compression methods allows their use in low SNR scenarios. The width of the obtained pulses is significantly smaller compared to the width of the LFM pulses. Having narrow pulses enables higher ToA and distance estimation resolution.

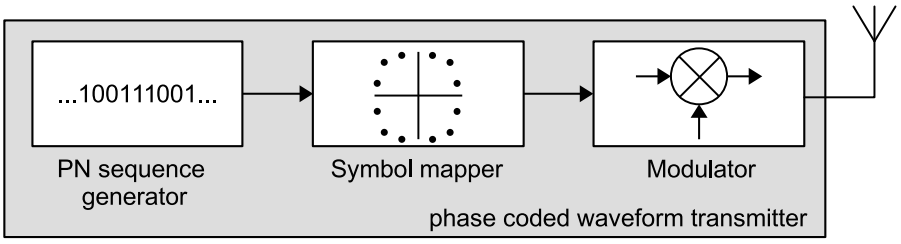
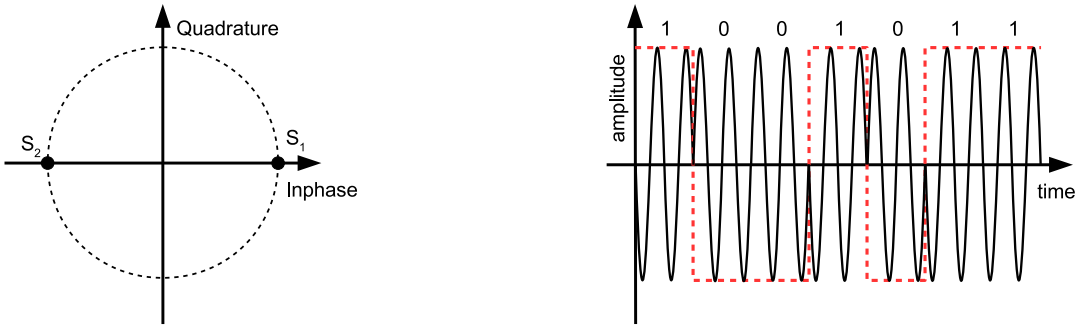


Fig. 3.10 A simplified block diagram of a phase coded waveform transmitter.



(a) BPSK symbols S1 (0 rad) and S2 (π rad) on a polar plot

(b) BPSK modulated waveform

Fig. 3.11 BPSK modulation (symbols and waveform)

3.3.4 Phase coded pulse compression waveforms

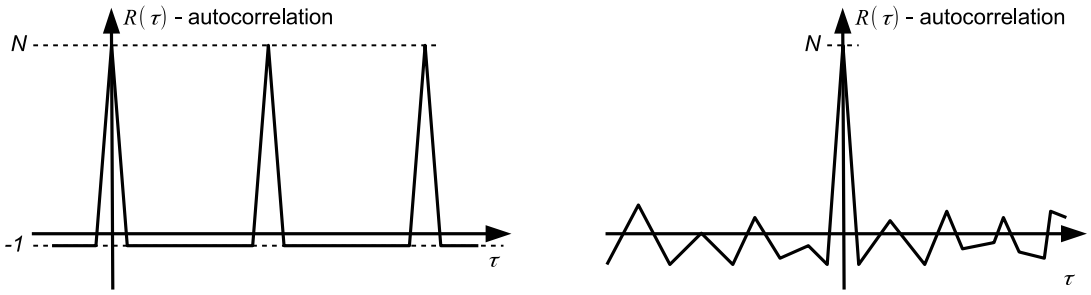
Phase coded waveforms are also often used in ranging and positioning systems. They consist of digital waveforms (bit or chip sequences), modulated using a phase modulation. This is a significant advantage in systems that combine ranging (i.e. positioning) and data communication, since the same, or a similar modulation scheme can be used for both functions.

A simplified block diagram of a system used for transmission of a phase coded waveform used for ranging is shown in Figure 3.10. Commonly used digital waveforms used for phase modulation of the carrier signal are binary[5] PN sequences. Binary Phase Shift Keying (BPSK) modulation is a common choice in ranging and positioning applications. The symbols used for BPSK modulation are shown in a constellation diagram in Figure 3.11a and a BPSK modulated digital waveform in Figure 3.11b.

The phase coded waveforms are used identically as the LFM pulses. The pulse acquired at the output of the matched filter is used for ToA estimation of the received phase coded waveform.

In ranging and positioning applications, linear PN sequences [39] are commonly used. They include: Maximum Length Sequences (MLS) [39], Gold codes [38], Kasami [60],

[5] Ternary or M-ary waveforms can also be used.



(a) Of a periodic PN sequence **(b)** Of an aperiodic PN sequence

Fig. 3.12 Autocorrelation function of a PN sequence

Walsh-Hadamard [4] etc. These sequences and their properties are further presented in this section. There are also nonlinear PN sequences [48], but they are of special interest mainly in cryptography.

Properties of the PN sequences

A few different properties of PN sequences are important in communication and ranging applications. They include: the **balance property**, **run-length distribution**, **autocorrelation**, **cross-correlation** and **frequency spectrum**.

The **balance property** represents the balance between the zeros and ones in the PN sequence. If the number of ones and zeros are not balanced, a DC component would exist in the baseband signal supplied to the modulator, eg. BPSK. This would limit the carrier suppression at the transmitter, since it strongly depends on the DC level present in the baseband signal.

A **"run"** is a subsequence of ones or zeros present in one PN sequence. It is usually desired that the runs of length one are half in the sequence, runs of length two are quarter of the sequence, runs of length 3 are one eighth of the sequence length, etc.

The **autocorrelation function** of the PN sequence is probably the most important property in ranging and positioning applications which use the pulse compression method. The autocorrelation function should have a narrow peak for $\tau = 0$ and should be equal to 0 for $\tau \neq 0$. For periodic PN sequences this can be almost (for $\tau \neq 0$ would be -1 instead of 0) achieved, as shown in Figure 3.12a, but for a single PN sequence the autocorrelation function absolute values for $\tau \neq 0$ would be significantly larger than 0, as in Figure 3.12b,. Periodic PN sequences are used in GPS, but in indoor localization systems, which use Time Division Duplex (TDD) multiplexing, aperiodic PN sequences are a common choice.

The **cross-correlation** function between different PN sequences of a same family is another important feature. This feature is essential in code division multiplexing systems. One example

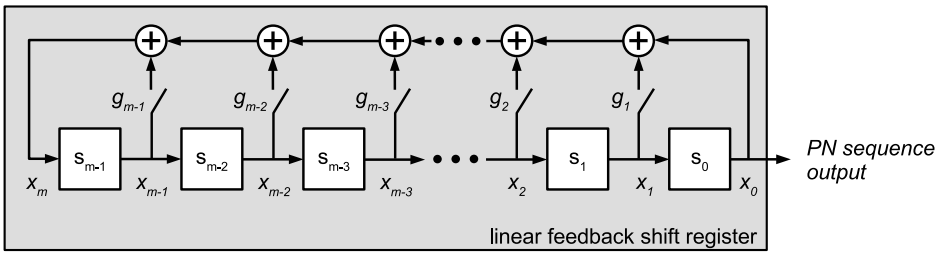


Fig. 3.13 Linear feedback shift register for generating a Maximum Length Sequences (MLS)

is the GPS system. Every satellite in the GPS system uses a different PN sequence. These PN sequences are almost orthogonal to each other, meaning that their cross-correlation functions are nearly zero. The signals from all the satellites can be received at the same time and decoded separately. This allows to simultaneously estimate the distance from the GPS receiver to multiple GPS satellites.

The **power spectrum** of a periodic PN sequence is represented by discrete lines in the frequency domain. On the other hand, the power spectrum of an aperiodic PN sequence is continuous and is calculated as a Fourier transform of its autocorrelation function [112]. The autocorrelation function of PN sequences used in ranging applications is, in an ideal case, a single narrow pulse. The spectrum of a single narrow pulse and, therefore, the spectrum of this PN sequence would be $\frac{\sin(x)}{x}$ or a $sinc(x)$ function.

Maximum Length Sequences (MLS)

The most common PN sequences considered in Direct Sequence Spread Spectrum (DSSS) systems as well as in ranging and pulse compression systems are the binary Maximum Length Sequences (MLS) or also referred to as m-sequences. These sequences are generated using linear shift registers as shown in Figure 3.13. The elements s_{m-1}, s_{m-2} to s_0 are 1-bit memory elements, forming the shift register. The outputs of each memory element, x_{m-1}, x_{m-2} to x_0, using the switches g_{m-1}, g_{m-2} to g_1, are *xor*-ed and fed back as an input of the shift register. The sequence is generated on the output x_0. The characteristic polynomial of a m-sequence is given by

$$f(x) = \sum_{i=0}^{m} g_i x^i \tag{3.21}$$

The degree of the polynomial is m (where $m > 0$). For $g_m = 1$ and $g_0 = 1$, the recurrence equation, describing the shift register in Figure 3.13, can be obtained as

$$x^m = g_{m-1}x^{m-1} \oplus g_{m-2}x^{m-2} \oplus \cdots \oplus g_1 x^1 \oplus g_0 \tag{3.22}$$

Table 3.2 Primitive polynomials for generating *m*-sequences with different lengths

shift register length m	sequence length $2^m - 1$	polynomial taps	# of primitive polynomials
2	3	[2, 1]	2
3	7	[3, 1]	2
4	15	[4, 1]	2
5	31	[5, 3], [5, 4, 3, 2] [5, 4, 2, 1]	6 6
6	63	[6, 1], [6, 5, 2, 1] [6, 5, 3, 2]	6
7	127	[7, 1], [7, 3], [7, 3, 2, 1], [7, 4, 3, 2] [7, 6, 4, 2], [7, 6, 3, 1], [7, 6, 5, 2] [7, 6, 5, 4, 2, 1], [7, 5, 4, 3, 2, 1]	18
8	255	[8, 4, 3, 2], [8, 6, 5, 3], [8, 6, 5, 2] [8, 5, 3, 1], [8, 6, 5, 1], [8, 7, 6, 1] [8, 7, 6, 5, 2, 1], [8, 6, 4, 3, 2, 1]	16

The sum in (3.22) is a modulo 2 sum (xor operation). The binary values $g_{m-1}, g_{m-2}, \ldots, g_0$, are the coefficients of the characteristic polynomial of the *m*-sequence.

Not every polynomial (determined with the coefficients g_i) can be used to generate *m*-sequences. They should be a *primitive polynomial* in order to generate MLS, i.e. *m*-sequence. The primitive polynomial ensures that the shift register (s_{m-1}, s_{m-2}, …, s_0), would cycle through all possible states, except the all zeros state.

Finding the primitive polynomials is a task involving mathematical methods which are out of the scope of this work. Nevertheless, a lot of work is available on the subject, as well as tables with computed primitive polynomials [22, 44, 62, 103]. In Table 3.2, primitive polynomials for sequences up to $m = 8$ are given. In the "polynomial taps" column, the coefficients $g_i = 1$ are given.

The *m*-sequences have good balance and run-length distribution. The periodic autocorrelation of the *m*-sequences is excellent. For a periodic sequence, the autocorrelation function looks like the one shown in Figure 3.12a. The peak value is equal to $2^m - 1$, and the rest of the values are equal to -1. Due to excellent autocorrelation properties, the *m*-sequences are perfect candidates for ranging and synchronization applications.

Table 3.3 Cross-correlation values of a Gold sequences

L	N_c	Normalized values for cross-correlation	Frequency of occurrence
Odd	2^L-1	$-1/N_c$	≈ 0.50
		$-(2^{(L+1)/2}+1)/N_c$	≈ 0.25
		$(2^{(L+1)/2}-1)/N_c$	≈ 0.25
Even not $k\cdot 4$	2^L-1	$-1/N_c$	≈ 0.75
		$-(2^{(L+1)/2}+1)/N_c$	≈ 0.125
		$(2^{(L+1)/2}-1)/N_c$	≈ 0.125

Nevertheless, *m*-sequences generated using different primitive polynomials, do not always have good cross-correlation properties. Therefore, they are rarely used in systems which use code division multiplexing (CDM).

Gold sequences

Another family of PN sequences, interesting for use in communication and localization applications, are the Gold sequences [38]. The main advantage of Gold sequences over *m*-sequences is the relatively low cross-correlation value between different sequences with same lengths. Therefore, the Gold sequences are very good candidates for use in CDM and code division multiple access (CDMA) systems. They are already used in the GPS system, where different satellites transmit data on the same frequency and at the same time.

The Gold sequences can have only three different values in the cross-correlation function. These values are given in Table 3.3. According to [79], the autocorrelation properties of the Gold sequences do not outperform those of *m*-sequences.

The Gold sequences are generated using two *m*-sequence generators. The outputs of these generators are summed (modulo 2) in order to generate a Gold sequence, as shown in Figure

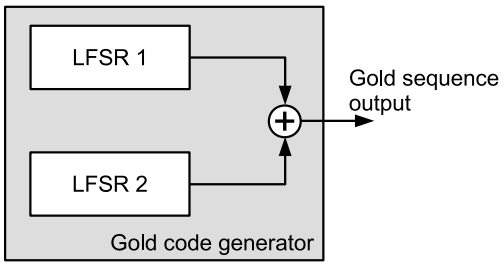


Fig. 3.14 Gold sequence generator

3.14. To generate a Gold sequence, a so called "preferred pair" of *m*-sequences must be used. Different Gold sequences are generated by changing the phase position of the two *m*-sequences. Therefore, a lot of different Gold sequences can be generated, using the same preferred pair. Assuming that registers of length L are used for generating the *m*-sequences, there can be $2^L - 1$ different phase positions of the two *m*-sequences. This leads to a total of $2^L - 1$ different Gold sequences plus the two initial *m*-sequences. Therefore, for a suitably large L, a large set of codes, with controlled cross-correlation value can be obtained.

Walsh-Hadamard sequences

Walsh-Hadamard sequences are one more family of PN sequences commonly used in data communication and positioning applications. A detailed description of these sequences can be found in [57]. These PN sequences are Walsh sequences, constructed using Hadamard matrices, therefore, the name Walsh-Hadamard. A Hadamard matrix H of order n is an *nxn* matrix, having only "1" and "-1" entries and satisfies $HH^T = nI_n$, where I_n is an identity matrix of order n. Hadamard matrices of order 1 and 2 are given by

$$H_1 = \begin{pmatrix} 1 \end{pmatrix} \tag{3.23}$$

$$H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \tag{3.24}$$

Hadamard matrices exist for orders which are multiples of 4. For generating Walsh sequences, Hadamard matrices of order 2^N are used. The form of the Hadamard matrices of order 2^N is given by

$$H_{2N} = \begin{pmatrix} H_N & H_N \\ H_N & -H_N \end{pmatrix} \tag{3.25}$$

The rows of the H_{2N} matrix form a set of Walsh-Hadamard sequences. These sequences are mutually orthogonal. In communication systems, in order to preserve the orthogonality between the sequences, they must be almost perfectly synchronized (aligned). The Walsh-Hadamard sequences are used in IS-95[6] systems for performing code division multiple access (CDMA).

Besides already mentioned linear PN sequences, there are a few more families of PN sequences which can be used for ranging and wireless data transmission purposes. Nevertheless, they all have very similar characteristics with respect to the already discussed PN sequences.

[6]The second generation of mobile communication standard in USA uses CDMA as an access technology. Its proprietary name is *cdmaOne*.

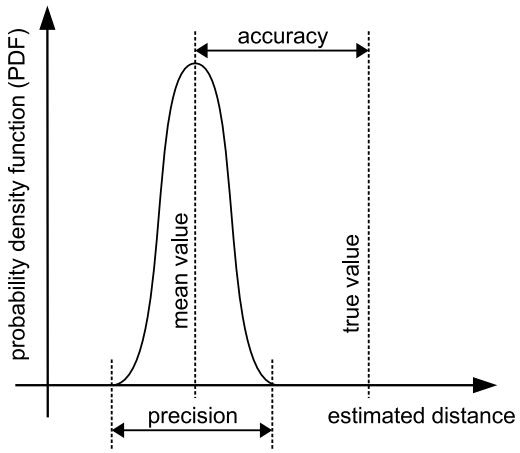


Fig. 3.15 Accuracy and precision

3.4 Ranging precision, accuracy and error sources

The distance estimates obtained using different ranging methods would always contain errors due to different factors. These estimates would follow a given probability distribution function as shown in Figure 3.15.

Precision is defined as a closeness of the obtained estimates, or measurements, under the same conditions. It is a measure for the quantity of random errors. Accuracy is a closeness of a given measurement, or a mean value of multiple measurements, to the true, or reference, value. It characterizes the systematic errors which can be easily compensated. The concepts of precision and accuracy are shown in Figure 3.15.

In this section, the most common systematic and random errors would be presented as well as the available approaches for reduction and compensation.

In ToF based ranging methods four main error sources would affect the ranging precision and accuracy. They include: **multipath propagation** of the RF signal, **noise**, **finite sample rate** of the A/D converters and the **clock frequency offset** between the wireless nodes.

Due to **multipath propagation**, several copies of the transmitted waveform would arrive at the receiver. In narrow band RF systems it would be impossible to resolve the time of arrival of the received waveform copies. Wide band systems, like UWB or 60 GHz systems, are more suitable for precise RF ranging since they can resolve the signals traveling different paths.

The thermal **noise** present at the receiver, can have a huge impact on the ranging precision. This is especially pronounced in RSS or power decay based ranging systems. With increase of distance between the nodes, the received power becomes lower, making the noise power dominant. In this case, the signal power cannot be used for precise distance estimation. In the

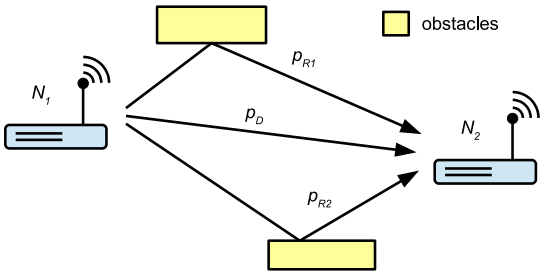


Fig. 3.16 Multipath propagation scenario. Multiple reflected signals arrive at the receiver.

case of ToF based methods, the situation is slightly different. The use of pulse compression methods increases the immunity to noise, at a price of a more complicated system.

The **finite sample rate** of the D/A and A/D data converters can also affect the ranging and positioning precision. According to Nyquist, sampling with $f_s = 2B$, where B is the signal bandwidth, is sufficient for perfect reconstruction of the signal. The reconstructed signal should be represented as a sum of $\frac{\sin x}{x}$ functions. These functions are infinite and, therefore, perfect reconstruction is not possible. Small errors would be present and would affect ranging precision.

The **clock frequency offset** of the clocks used for ToF measurement can bias the range and position estimates. The commonly used crystal oscillator clock sources have excellent frequency accuracy and drift (10 - 50 ppm), which makes them suitable for time of flight measurement. The problem appears in TWR scenarios, where the reply time is significantly larger compared to the the time of flight of the transmitted waveform. In this case, the time estimation error, due to the large reply time, can be comparable or even larger than the time of flight. This error can be reduced by using a clock generator with higher accuracy and precision. This is usually not possible in hand held devices and wireless sensor networks, due to the robustness of these clock generators or due to their price.

3.4.1 Multipath propagation influence on ranging precision

The multipath propagation can be significantly pronounced in indoor environments due to multiple obstacles present. A simple multipath propagation scenario is shown in Figure 3.16. The waveform traveling the direct (shortest) path, p_D, is called direct path (DP) wave, or line-of-sight (LOS) wave. In ranging and positioning applications, the estimation of the time of arrival of the DP is of interest. Nevertheless, a waveform being reflected from different obstacles can interfere with the DP waveform at the receiver. The estimation of the time of arrival of the DP waveform would be challenging in this case.

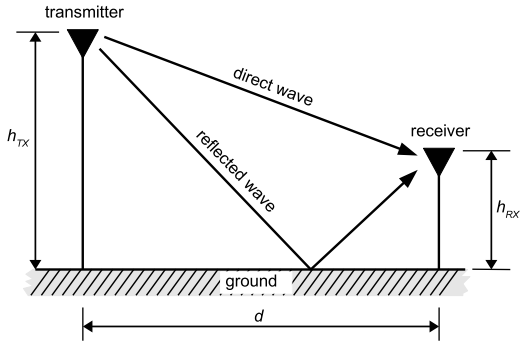


Fig. 3.17 Two-Ray ground-reflection diagram used for derivation of the channel model

The simplest multipath propagation model is the Two Ray Model [91]. This model takes into account the direct wave as well as the ground reflected wave, as shown in Figure 3.17. In this model, the transmitted waveform travels to the receiver using two different paths. These two paths have different lengths, leading to two different times of flight. If a precise propagation model is required, more than two paths should be included. The Ray tracing propagation model [34] includes significantly more than two paths and is commonly used.

In indoor environments strong multipath with many reflections is present [99]. Using all of the reflections for channel modeling is not the optimal solution. The overall effect of these reflections on the channel impulse response is of interest and, therefore, statistical channel models are commonly used. The most common statistical model for indoor multipath propagation is the Saleh-Valenzuela channel model [95]. For millimeter wave (mmWave) systems, used indoors, the ground reflection plays a significant role and, therefore, the statistics of the reflection are combined with the SV model to form the TSV channel model described in [97].

Channel impulse response

The wireless channel can be modeled as Linear Time Invariant (LTI) system as

$$h(t) = \sum_{i=0}^{\infty} \beta_i e^{j\Phi_i} \delta(t - \tau_i) \tag{3.26}$$

The impulse response of the channel model in (3.26) consists of the direct wave and multiple reflected waves, summed together. Every reflection is amplified (attenuated) by β_i. The angle Φ_i represents the phase shift introduced in every separate path and the τ_i is the delay of each signal. The $\delta(\cdot)$ is the Dirac delta function.

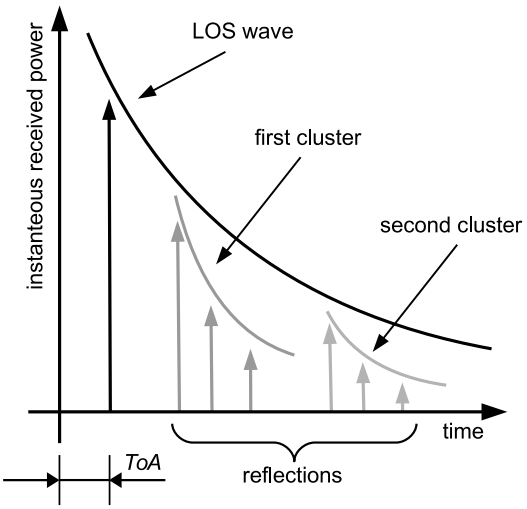


Fig. 3.18 Indoor channel impulse response (according to SV channel model)

According to [95], the reflections are grouped in clusters, and their power decays exponentially. Also, the overall signal power decays exponentially, as can be seen in Figure 3.18. The TSV model impulse response is quite similar to the one shown in Figure 3.18.

Since the reflections are grouped in clusters, in some cases it can be better to represent the channel model from (3.26) as a sum of clusters, which are sums of pulses, as

$$h(t) = \sum_{l=0}^{\infty} \sum_{k=0}^{\infty} \beta_{kl} e^{j\Phi_{kl}} \delta(t - T_l - \tau_k) \tag{3.27}$$

In (3.27), T_l is the start time of the l-th cluster. The rest of the symbols are the same as in (3.26).

Using measurements made in [95], statistical parameters are extracted for different scenarios, shown in Table 3.4. Using the parameters for the given channel models, new sets of channel impulse responses can be generated.

Time of arrival estimation in the presence of multipath propagation

ToA estimation in a multipath propagation environment can be a challenging task. For a single transmitted pulse, multiple pulses would be received at the receiver due to multipath propagation. In indoor scenarios, it can occur that the transmitted pulse, $s(t)$, is wider than the difference of the arrival delays, $\tau_i - \tau_{i+1}$, from (3.26). This would lead to overlapping of the pulse traveling the LOS path with the pulse being reflected from an obstacle. The result is shown in Figure 3.19. In this case, the leading edge of the first received pulse is not affected significantly. A more complex case is shown in Figure 3.20 where the difference in length of the direct path and the indirect path is small compared to the pulse length. The leading edge

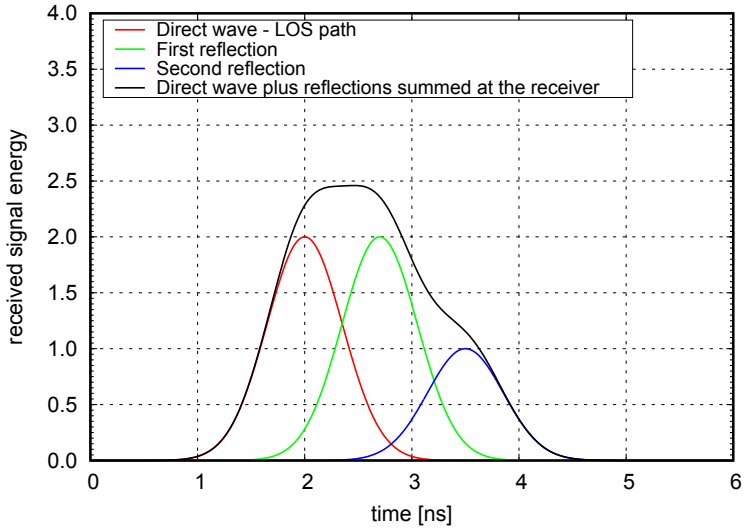


Fig. 3.19 Slightly overlapping multiple received pulses in a multipath propagation environment

of the received waveform and the leading edge of the pulse traveling the LOS path are not overlapping. The peaks of the first two arriving pulses (the peak traveling the LOS path and the reflected one), are merged into one peak in the received waveform. In this scenario it is almost impossible to perform precise ToF estimation, since the start of the first pulse cannot be efficiently estimated.

If pulse compressed waveforms are used, the time of arrival estimation is performed, after matched filtering, i.e. correlation. The matched filter would produce a pulse per received waveform (direct path and reflected). These pulses can also overlap if multipath propagation is present.

Table 3.4 Scenarios and environments for the available SV and TSV models

Model name	Scenario	Environment
CM1	LOS	Residential
CM2	NLOS	Residential
CM3	LOS	Office
CM4	NLOS	Office
CM5	LOS	Library
CM6	NLOS	Library
CM7	LOS	Desktop
CM8	NLOS	Desktop
CM9	LOS	Kiosk

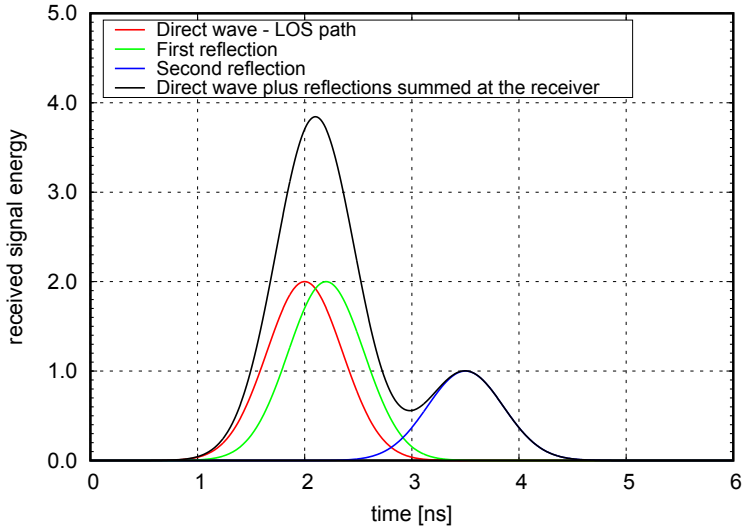


Fig. 3.20 Significantly overlapping multiple received pulses in a multipath propagation environment

Avoiding errors due to multipath propagation

In order to avoid ranging errors due to multipath propagation, a large channel bandwidth can be used. The UWB radio systems [3, 23, 65, 88, 104, 117] use the spectrum between 3.1 GHz and 10.6 GHz. The channel bandwidth can reach up to several hundred MHz. In ranging applications, the impulse radio UWB (IR-UWB) system is commonly used. It is a carrierless system which transmits very short pulses [7] and, therefore, providing excellent spatial resolution, usually better than 10 *cm*. The 60 GHz based systems also use large channel bandwidths of up to 2 GHz. A spatial resolution of down to a few *cm* can be achieved [25, 73, 83]. The large channel bandwidth allows to easily identify the LOS component and the reflected components of the signal arriving at the receiver. Identifying the LOS component is essential in strong multipath propagation environments, e.g. indoors.

The second approach for minimizing the multipath propagation influence on ranging precision is to use beamforming or directional antennas. This can be beneficial especially in narrowband systems, where identification of the direct and reflected components of the signal arriving at the receiver is not possible. The use of beamforming or directional antennas would filter out the reflection arriving from directions other than the direction where the transmitter is. This approach is actively used, especially in local and indoor positioning system [26, 94, 106].

In narrowband channels, where no beamforming antennas are available, super-resolution methods are often used. These methods [50, 51, 55, 63, 70] utilize super-resolution methods as

[7] The pulses have duration of under 1 ns and in the most cases a few hundreds of picoseconds, depending of the channel bandwidth.

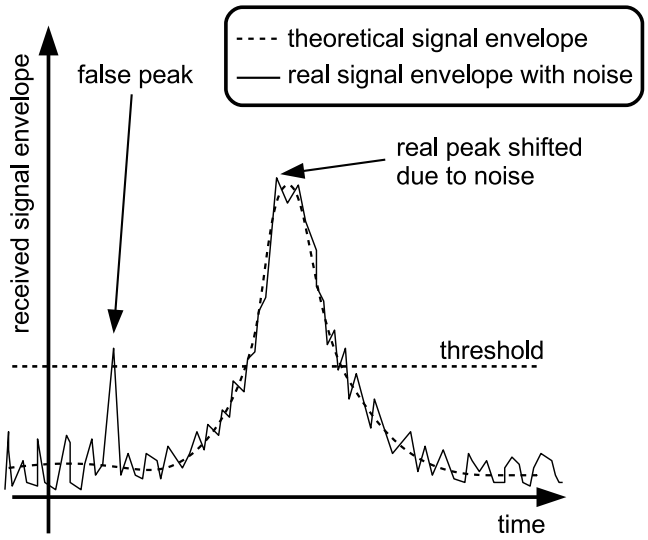


Fig. 3.21 False detection of the first arrival pulse and errors in ToA estimation

Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT), Root-Multiple Signal Classification (RootMUSIC), Matrix Pencil (MP) or slightly modified versions of them.

3.4.2 Noise and channel bandwidth influence on ranging precision

Noise and limited channel bandwidth can significantly affect precision of the time of arrival estimation of the received waveform. The time of arrival estimation is basically a pulse position estimation problem. This pulse can be received by the receiver, e.g. UWB, or it can be obtained using pulse compression methods.

In Figure 3.21 a single received pulse, with additional Gaussian noise is shown. The time of arrival, i.e. position estimation of this pulse must be performed in order to further estimate the distance between the nodes. Two main effects are responsible for the errors in pulse position, i.e. ToF estimation. It might occur that the noise at a certain moment achieves a large value, exceeding the threshold level at the receiver. A false peak, shown in Figure 3.21 would be detected and a large ranging error would be observed.

Another effect appears when a pulse is received and the noise added to the signal shifts the edges of this pulse. The leading edges have finite rise time due to the limited bandwidth. The peak would be generally detected, but precise detection of the edge, i.e. peak position, would not be possible. If infinite channel bandwidth is assumed, the received pulse would have sharp leading and falling edges and would be immune to the effect of the added noise.

The ToA estimation error due to noise and finite channel bandwidth can be bounded. The lower bound of the variance of a time of arrival estimator (i.e. distance estimator), shown in

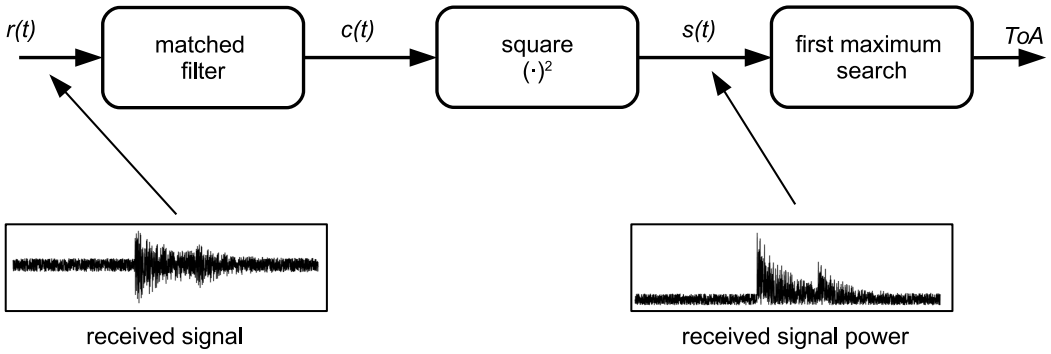


Fig. 3.22 Simple time of arrival estimator

Figure 3.22, is given with the Cramér-Rao Lower Bound (CRLB). This bound is important for comparing different ranging as well as positioning methods.

Cramér-Rao lower bound

The Cramér-Rao Lower Bound (CRLB) is used to bound the theoretical minimum of the variance of an unbiased estimator [20, 89, 90]. It is also used as a lower bound for the variance of distance estimators. This lower bound usually cannot be achieved in practical implementations, but can be used for comparison of different estimators.

In the general case, it can be assumed that θ is a deterministic parameter to be estimated, given a set of noisy measurements, x. These measurements are distributed according to some probability density function $f(x;\theta)$.

The CRLB is bounding the variance $var(\hat{\theta})$, of the estimate, $\hat{\theta}$. The estimate, $\hat{\theta}$, represents an estimation of the unknown deterministic parameter θ. An unbiased estimator is an estimator for which $E[\hat{\theta}] - \theta = 0$, where $E[\cdot]$ represents the expected value. The CRLB does not guarantee the existence of such an estimator that is able to obtain the bound. Nevertheless, if an estimator achieving the CRLB exists, then it would be a Minimum-Variance Unbiased Estimator (MVUE) [61]. The CRLB is not tied to a specific estimator, i.e. it is general no matter which estimator is used.

In a general case, if the probability density function $f(x;\theta)$ satisfies

$$E\left[\frac{\partial \ln f(x;\theta)}{\partial \theta}\right] = 0 \tag{3.28}$$

the variance, $var(\hat{\theta})$, of $\hat{\theta}$ can be bounded with

$$var(\hat{\theta}) \geq \frac{1}{-E\left[\frac{\partial^2 \ln f(x;\theta)}{\partial \theta^2}\right]} = \frac{1}{E\left[\left(\frac{\partial \ln f(x;\theta)}{\partial \theta}\right)^2\right]} = \frac{1}{I(\theta)} \tag{3.29}$$

The derivative is evaluated at the true value of θ. The expectation (E) is with respect to $f(x;\theta)$. The function $I(\theta)$ is the Fisher information [68].

The CRLB can be also used to bound the variance of an estimator used for range estimation. The details of derivation of the CRLB of the range estimates can be found in [111]. The CRLB of the range estimates can be calculated for two separate cases. The first case is the simple case, where only the direct wave is present, i.e. no reflections arrive at the receiver. The signal model in this case can be represented as

$$r(t) = \beta e^{j\varphi} s(t-\tau) + w(t) = \underline{\beta} s(t-\tau) + w(t) \tag{3.30}$$

where the attenuation, $\underline{\beta} = \beta e^{j\varphi}$, is complex. The function $s(t)$ represents the shape of the received pulse. The delay, τ, is the time of flight. The CRLB in this case is given by

$$var(\hat{r}) = \sigma_{\hat{r}}^2 \geq \frac{c^2}{(2\pi B)^2 E_s/N_0} \tag{3.31}$$

where $\hat{r}$ is the distance estimate, c is the speed of light, B is the bandwidth of the pulse and E_s/N_0 is the signal energy to noise density ratio. The bandwidth B represents the RMS bandwidth of the pulse used for ranging. It is calculated as

$$B = \sqrt{\frac{\int_{-\infty}^{\infty} (2\pi f)^2 |S(f)|^2 df}{\int_{-\infty}^{\infty} |S(f)|^2 df}} \tag{3.32}$$

where $S(f) = \mathscr{F}\{s(t)\}$. The signal power is $P_s = E_s/t_s$, where t_s is the duration of the pulse or the ranging waveform. The power of the noise is $P_n = N_0 B$. Finally, the signal energy to noise density ratio can be expressed as $E_s/N_0 = t_s B \cdot SNR$.

When a single pulse is transmitted, the time-bandwidth product would be $t_s B = 1$. A waveform n times longer would have $t_s B = n$. Usually, a PN sequence with length n is preferred over a single pulse. The CRLB would be reduced if a waveform with larger time bandwidth product is used.

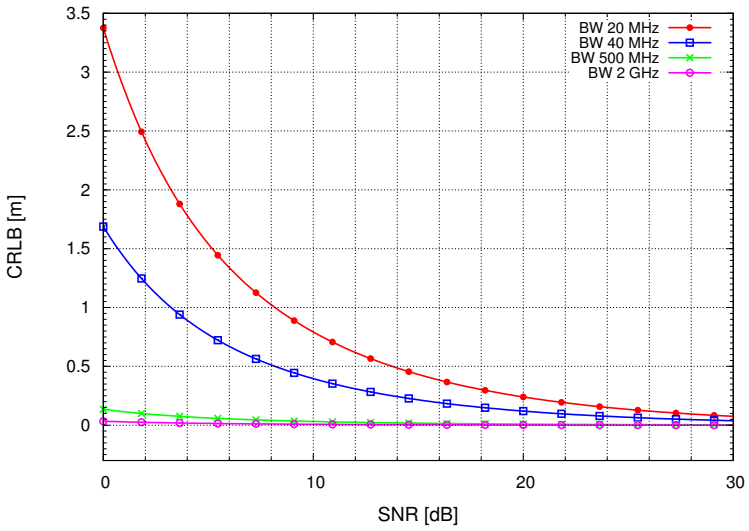


Fig. 3.23 CRLB as a function of channel bandwidth and SNR for a single received pulse ($\tau B = 1$)

For a multipath propagation environment the CRLB would slightly differ compared to (3.31). Using the channel impulse response form (3.26) and adding additive noise, $w(t)$, the received signal model would be

$$r(t) = \sum_{i=0}^{\infty} \beta_i e^{j\varphi_i} s(t-\tau_i) + w(t) \tag{3.33}$$

The CRLB of the distance estimate, performed using the signal model in (3.33), is given by

$$var(\hat{r}) = \sigma_{\hat{r}}^2 \geq \frac{c^2}{(2\pi B)^2 E_s/N_0} \left(1 + \frac{1}{E_s/N_0}\right) \tag{3.34}$$

It can be noticed that the CRLB is slightly higher, by a factor of $1 + \frac{1}{E_s/N_0}$, in this case. The details about derivations of (3.34) can be found in [111]. For high values of E_s/N_0, (3.34) converges to (3.31).

Figure 3.23 plots the CRLB as a function of SNR of the received signal. The CRLB is plotted for $t_s B = 1$ and for channel bandwidths of 20, 40, 500 and 1000 MHz.

It is widely known that in nonlinear estimation problems, as the one which results are shown in Figure 3.22, the performance of the estimator is characterized by different SNR regions. The CRLB is commonly used as a performance benchmark of an estimator, but it is well known that for medium and low SNR it is inaccurate. This is due to the so called threshold effect, studied in [52, 69, 77, 115]. A more accurate lower bound for the medium and low SNR regions is the Ziv-Zakai bound (ZZB) [120], or its improved versions such as the Bellini-Tartara

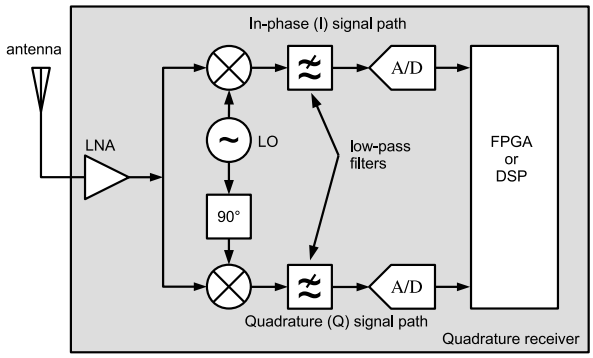


Fig. 3.24 Simplified block diagram of a quadrature receiver

bound [6], the Chazan-Zakai-Ziv [15] bound and the Weiss-Weinstein bound [115]. The main disadvantage of the Ziv-Zakai bound is that it is more complex for evaluation.

3.4.3 Finite sampling rate impact on ToF estimation precision

Another factor affecting the precision of the ToF estimation in ranging and localization systems is the finite sampling rate of the A/D converters. A typical quadrature receiver used for reception of phase coded waveforms is shown in Figure 3.24. The received signal is first demodulated and sampled by the A/D converters. The sampling rate of these converters is, in the worst case, a couple of times higher than the Nyquist rate. The mixers used for down-converting the signal and the low-pass filters are implemented as analog parts.

The Nyquist theorem states that a waveform sampled with a sample rate of $f_s = 2B_s$, where f_s is the sampling rate and the B_s is the signal bandwidth, can be perfectly reconstructed. The reconstruction is performed by introducing *sinc* functions where the samples were originally sampled. In real life applications there are a few issues preventing ideal reconstruction of the sampled signal. The first issue is the quantization of the acquired waveform. The sampled waveform is quantized in a finite number of discrete levels by rounding or truncating. Therefore, the discrete samples would not represent the true value of the signal at the moment of sampling. This effect is similar to adding additional noise to the waveform, and is therefore quantified with a signal to quantization noise ratio (SQNR) parameter given by

$$SQNR = 20\log_{10}\left(2^{Q}\right) \approx 6.02 \cdot Q \ [dB] \tag{3.35}$$

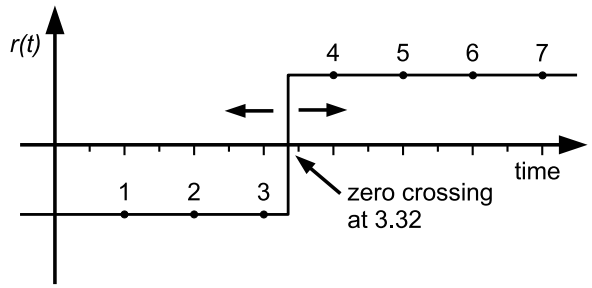


Fig. 3.25 A total of seven samples acquired by sampling a step pulse

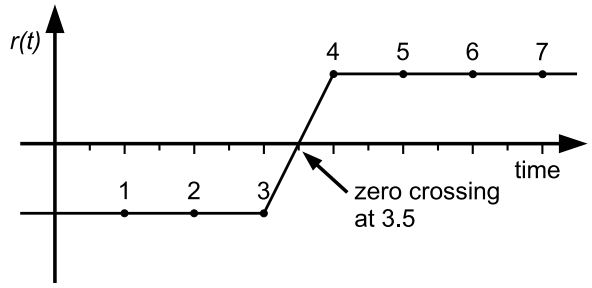


Fig. 3.26 First order (linear) interpolation between the samples of the step function

The second issue is the reconstruction of the waveform using *sinc* functions. The reconstructed waveform $r_r(t)$ would be

$$r_r(t) = \sum_{-\infty}^{\infty} s(nT)sinc(t - nT) \tag{3.36}$$

where $s(nT)$ are the samples taken in time instances nT, and $T = 1/f_s$ is the sampling period. It can be clearly seen that (3.36) is unusable in practical applications due to the infinite length of the *sinc* function.

In ToF based ranging, the precise detection of the time of arrival of the signal is crucial. A simple example is shown in Figure 3.25. A step pulse is transmitted and received by the receiver. This step pulse is sampled and represented with the samples 1 to 7. The zero-crossing point, at position 3.32 on the time axis marked with an arrow, is the true zero-crossing point. However, given the sampling instants 1 to 7, this point only can be estimated at either position 3 or 4. This is the simplest approach and is equivalent to zeroth order interpolation[8]. An approach with higher order interpolation is shown in Figure 3.26. The interpolation between the samples is linear, i.e. of first order. In this case, the estimated zero crossing position does not have to be an integer. This is a rather simple case and probably does not illustrate completely the

[8]The zeroth order interpolation can be implemented using zero order hold.

advantage of using a larger interpolation order. Nevertheless, a larger order interpolation would increase the signal reconstruction precision, leading to smaller timing errors.

If no interpolation is performed, the range estimation error is uniformly distributed in a range bin. For example, all ranges between 3 and 4 would be estimated as range 3. This is called range binning and the ToF error associated with the range binning can be calculated as

$$\sigma^2_{sample} = \frac{1}{12 f_s^2} \tag{3.37}$$

The range binning error can be reduced by increasing the sampling frequency which is justified only if the range binning error is dominant. Nevertheless, this is a costly method and, therefore, usually avoided.

The simplest approach for reducing the error due to range binning is to interpolate between the acquired samples. In [83], a Fast Fourier Transform (FFT) zero padding interpolation is used. Other interpolation methods are described and used for ToA estimation in [27].

A novel Compressive Sensing (CS) ToA estimation algorithm which reduces the ranging errors due to range binning is proposed in [118]. They achieve higher ToA estimation precision using lower sampling rate, compared to the other CS based ToA estimation algorithms, like the one in [35].

Super-resolution methods are also used for precise ToA estimation in scenarios with finite sample rate of the A/D converters. The super-resolution method described in [31] uses a modified version of the MUSIC algorithm and achieves a subsample ToA estimation precision. Nevertheless, the super-resolution methods, as well as the compressive sensing based methods are significantly complex.

The GPS receivers use oversampling A/D converters, but also uses an approach for additional oversampling of the received signal. This approach is depicted in Figure 3.27. The GPS has a sample rate higher than the Nyquist rate of the received signal. At the same time, the sampling rate of the A/D converter is not a multiple of the symbol rate of the incoming BPSK modulated PN sequence. The benefit of having a non-integer ratio between the symbol rate and the sample rate is that every successive symbol would be sampled at different time instants, with respect to the symbol interval. Since the GPS satellites transmit the same PN sequence periodically, two successive sequences would not be sampled at the same time instants with respect to the start of the sequences. Therefore, after a significantly long period of time, an oversampled version of the PN sequence would be obtained. The details of the approach are given in [108].

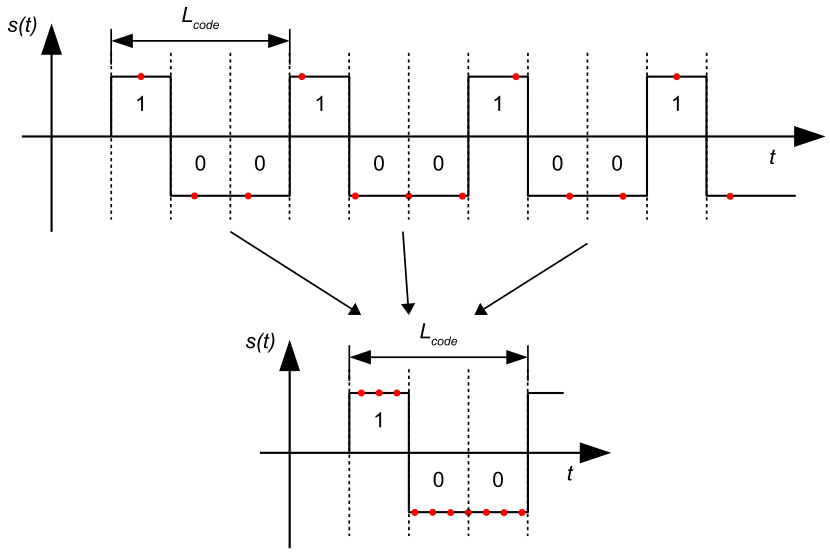


Fig. 3.27 Oversampling approach used in the GPS system

This approach is illustrated in Figure 3.27. Three copies of a periodic PN sequence (in this case *[1,0,0]*), are received. The sampling rate is fixed and the acquired samples are marked with red dots. It can be noticed that the three copies of the PN sequence are not sampled at the same time instants.

In this work, a method for reduction of the range binning error is presented. This method achieves higher effective sampling rate at the cost of additional signal processing complexity of the ranging waveform, as well as, increased wireless medium usage. The sample rate of the A/D converter is not increased in the proposed approach. A detailed description is given in Chapter 4.

3.4.4 Clock frequency accuracy

Clock frequency accuracy directly affects the ToA estimation accuracy. Clock generators in electronic devices use quartz crystal oscillators, due to their good characteristics regarding frequency stability and accuracy. Nevertheless, due to manufacturing imperfections, quartz crystals have an initial frequency offset, to which aging and temperature variations could add up additional frequency drift. The initial frequency offset is usually in the order of $\pm 0.25\ ppm$ to $\pm 100\ ppm$ [9]. The frequency drift due to temperature change and aging is in the range $\pm 10\ ppm$ to $\pm 50\ ppm$ for commercial and industrial temperature ranges[10]. If better temperature stability

[9] $1\ ppm = 1 \cdot 10^{-6}$

[10] Commercial temperature range is $0 \div 70^{\circ}C$ and industrial temperature range is $-40 \div 85^{\circ}C$ [78].

is required, temperature compensated crystal oscillators (TCXO) and oven compensated crystal oscillators (OCXO) can be used. Their temperature stability is less than a few ppm.

In commercial electronic devices and in wireless sensor networks, quartz crystal oscillators without any temperature compensation are a common choice. For applications like synchronizing a telecommunication network, or GNSS applications, atomic precision clocks are often used. The frequency stability of uncompensated quartz crystal oscillators is sufficient for performing precise time and range measurements.

The quartz oscillator frequency accuracy, σ_f, is defined as

$$\sigma_f = \frac{f_{true} - f_{nom}}{f_{nom}} \tag{3.38}$$

where the frequency f_{true} is the true frequency on which the oscillator is oscillating, and f_{nom} is the nominal frequency.

Since the time is measured by counting periods of a certain clock signal, a frequency inaccuracy of $\pm\sigma_f$ would introduce a time measurement error of $\mp\sigma_f$. The clock frequency inaccuracy can significantly affect the ToF estimation and introduce large ranging error, depending on the used ranging method.

Errors in ToA based ranging

In a time of arrival approach shown in Figure 3.2a, the node N_1 transmits a waveform at time t_{tx}, in order to perform ToF estimation. The node N_2 receives the same waveform at time t_{rx}, after t_{tof}. The transmission time instant, t_{tx}, is known and is negotiated between the nodes in advance. Node B calculates the time of flight as $t_{tof} = t_{tx} - t_{rx}$. Due to the clock frequency offset between the nodes they would not stay perfectly synchronized and should be synchronized periodically, before the ranging is performed. In Figure 3.2a the synchronization occurs at the time instant t_{sync}. After the synchronization, a time of t_{sm} seconds elapses before the ranging is performed.

In a real scenario clocks at nodes N_1 and N_2 would have frequency errors of σ_1 and σ_2 respectively. Therefore, the time measured by the nodes would contain an error directly proportional to the frequency errors. This also applies for the time from the synchronization event to transmission and reception of the ranging waveform. The time of flight, with the clock frequency error included, would be estimated as

$$\hat{t}_{tof} = t_{rx} - t_{rx} = \left[t_{sync} + (t_{sm} + t_{tof})(1+\sigma_2)\right] - \left[t_{sync} + t_{sm}(1+\sigma_1)\right] \tag{3.39}$$

where the $\hat{t}_{tof}$ is the estimated time of flight and the times t_{tx} and t_{rx} are the measured transmit and receive time instants, respectively. By rearranging equation (3.39), the estimated time of flight would be

$$\hat{t}_{tof} = t_{sm}(\sigma_2 - \sigma_1) + t_{tof}(1 + \sigma_2) = t_{sm}\sigma_{21} + t_{tof} \tag{3.40}$$

where $\delta_{21} = \delta_2 - \delta_1$ is the frequency offset between the clocks in the nodes N_2 and N_1. The product $t_{sm}\delta_{BA}$ is the ToF estimation error since the approximation $t_{tof}(1+\sigma_2) \approx t_{tof}$ can be performed. If the synchronization is performed long before transmission of the ranging frame, and the clock frequency offset between the two nodes is large, a significant ranging error can be expected. For example, assuming that the $t_{sm} = 1\ ms$ and the clock frequency errors are $\sigma_1 = -40\ ppm$ and $\sigma_2 = 40\ ppm$ [11], the time of flight estimation error would be $\varepsilon = t_{sm}\delta_{21} = 80\ ns$. This error, multiplied by c would produce a ranging error of $\varepsilon_r = \varepsilon \cdot c = 24\ m$.

Errors in TWR based ranging

For indoor ranging and localization, TWR based methods are usually preferred. These methods are also not immune to clock frequency offsets present in the nodes. The clock frequency offset sensitivity in this case is a direct consequence of the long replay time.

A typical TWR scenario is shown in Figure 3.2b. The node N_1 transmits a frame to the node N_2 at time t_{tx1}. Node N_2 receives this frame at time t_{rx2} and replies after t_{reply}, at time t_{tx2}. Finally, at time t_{rx1}, the frame is received back to node N_1. Node N_1 notes the time instants t_{tx1} and t_{rx1}. The duration t_{reply} is either negotiated in advance, or measured by the node N_2 and sent back to node N_1.

The time of flight can be estimated as:

$$\hat{t}_{tof} = \frac{(t_{tx1} - t_{rx1}) - (t_{rx2} - t_{tx2})}{2} \tag{3.41}$$

where the term $t_{tx1} - t_{rx1}$ is the round trip time (RTT) time and the reply time is $t_{reply} = t_{rx2} - t_{tx2}$. Since the round trip time in node N_1 is measured with a clock having frequency error of σ_1 and the reply time at node N_2 is measured with a clock having frequency error of σ_2, the equation (3.41) would become

$$\hat{t}_{tof} = \frac{(t_{tx1} - t_{rx1})(1 + \sigma_1) - (t_{rx2} - t_{tx2})(1 + \sigma_2)}{2} \tag{3.42}$$

[11] The clock inaccuracy specified here is in the order of $\pm 10 \div \pm 50\ ppm$ which is typical for temperature uncompensated quartz crystal clock generators. The value for t_{reply} of 1 ms, is also a realistic approximation.

The term $t_{tx1} - t_{rx1}$ can be substituted with $2t_{tof} + t_{reply}$, and the term $t_{rx2} - t_{tx2}$ can be substituted with t_{reply} and, therefore, the ToF estimate would be

$$\hat{t}_{tof} = \frac{(2t_{tof} + t_{reply})(1+\sigma_1) - t_{reply}(1+\sigma_2)}{2} \tag{3.43}$$

By rearranging (3.43), the final equation would be

$$\hat{t}_{tof} = t_{tof}\,(1+\sigma_1) + \frac{1}{2} t_{reply}\,(\sigma_1 - \sigma_2) \tag{3.44}$$

Two ToF estimation errors are present in (3.44). The first error is due to the direct impact of the clock frequency error on the ToF. This error is included in the first part of (3.44) and is equal to $t_{tof}\sigma_1$. Nevertheless, this error is reasonably small compared to the t_{tof} and, therefore, the approximation $t_{tof}\,(1+\sigma_1) \approx t_{tof}$ in (3.44) can be performed. The second error in ToF estimation is a result the clock frequency offset between the two nodes. The term $\frac{1}{2} t_{reply}\,(\sigma_1 - \sigma_2)$ in (3.44) represents this error. If both clock frequency errors σ_1 and σ_2 are equal, this error would be nullified. Usually, this is not the case in real scenarios and this error is significant and not negligible, mainly due to the large t_{reply}. The final expression for the ToF estimation is then given by

$$\hat{t}_{tof} = t_{tof} + \frac{1}{2} t_{reply}\,(\sigma_1 - \sigma_2) \tag{3.45}$$

Errors in cooperative ranging methods

In cooperative ranging methods [85] like N-Way Ranging (NWR) [40], the distance errors due to clock frequency offset between the nodes can become significantly large. The errors accumulate and, in indoor scenarios, they can be larger than the measured distances.

In Chapter 5, a compensation method is proposed. This method corrects the errors due to clock frequency offset and does not increase the system complexity.

Compensation of the ranging error due to clock frequency offset

In ToA based positioning, e.g. GPS, use of high accuracy clock sources is necessary. Nevertheless, different compensation methods are used in order to further improve the ranging precision and accuracy. The methods for clock frequency inaccuracy compensation are listed in [58].

The ToA based indoor positioning is avoided due to its complexity. Nevertheless, a method for keeping the positioning error due to clock frequency inaccuracy at minimum is presented in [101]. Using this method the transmissions used for synchronization are used also for

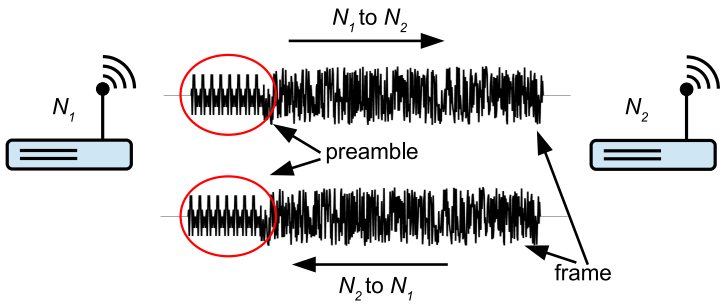


Fig. 3.28 Preamble (in red circle) used for clock frequency offset estimation in TWR

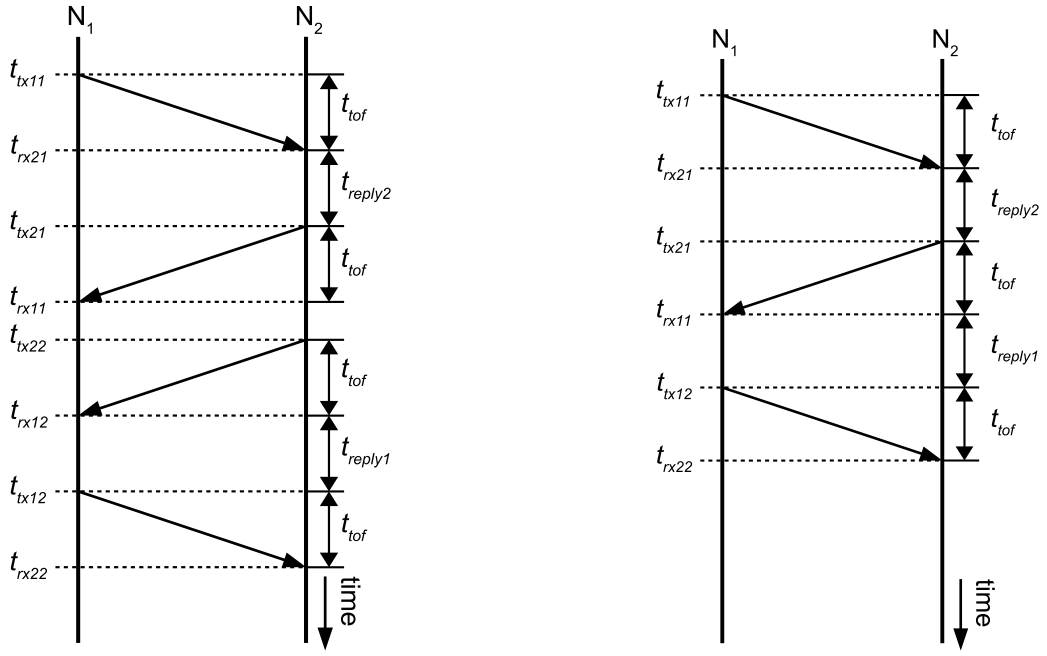
positioning of the nodes. The time between synchronization and positioning, in this method, is kept at a minimum, which leads to minimal positioning and ranging errors.

In TWR ranging scenarios two approaches for compensation of the ToF measurement are commonly used. The first approach is to estimate the clock frequency offset between the nodes and to correct the ranging measurement. In Figure 3.28, a typical scenario with clock frequency offset estimation is shown. The transmitted frames contain a preamble (in the red circle on Figure 3.28), mainly used for detection of the frame arrival at the distant node. This preamble, however, can be also used for estimating the frequency offset. When node N_2 receives a frame from node N_1 it estimates the clock frequency offset between the nodes. The same applies for node N_1 when a frame from node N_2 is received. Further details of this method are presented in [13].

It is assumed that the carrier frequency at the nodes N_1 and N_2 is obtained using the same clock frequency used for time measurement. In most of the devices this is the case, since only one crystal clock frequency source is used. If this is not the case, the symbol rate can be used for the same purpose. The symbol rate and timing are estimated at the receiver in order to perform decoding of the received signal. The symbol rate is usually derived from the clock source and, therefore, both of them would have the same inaccuracy.

Another approach for clock frequency offset error compensation in a TWR scenario is the Symmetrical Double Sided - Two Way Ranging (SDS-TWR) [43]. In this approach, the TWR is performed twice. A detailed scenario is shown on Figure 3.29a. Node N_1 initiates the TWR process by sending a frame to node N_2. After the corresponding reply is transmitted by node N_2 and received by N_1, the ToF is estimated as

$$\hat{t}_{tof1} = t_{tof} + \frac{1}{2} t_{reply} (\sigma_1 - \sigma_2) \tag{3.46}$$



(a) SDS-TWR using four transmissions. **(b)** SDS-TWR using three transmissions.

Fig. 3.29 Symmetrical Double Sided - Two Way Ranging (SDS-TWR) time diagram

One more TWR is performed, this time initiated by node N_2. The ToF estimate in this case would be

$$\hat{t}_{tof2} = t_{tof} + \frac{1}{2} t_{reply} (\sigma_2 - \sigma_1) \tag{3.47}$$

As can be seen from (3.46) and (3.47), the difference is in the sign in front of the term $\frac{1}{2} t_{reply} (\sigma_1 - \sigma_2)$. Finally, the values estimated in both nodes are exchanged and an average is calculated as

$$\hat{t}_{tof} = \frac{t_{tof1} + t_{tof2}}{2} = t_{tof} + \frac{1}{2} t_{reply} (\sigma_2 - \sigma_1) + t_{tof} + \frac{1}{2} t_{reply} (\sigma_1 - \sigma_2) = t_{tof} \tag{3.48}$$

According to (3.48), the estimated value would not include the error due to the clock frequency offset between the nodes. The main assumption in the described case is that both reply times are the same.

The ranging error due to clock frequency offset in (3.48) is fully compensated. The main problem in the SDS-TWR method is that it requires additional transmissions. In the case shown in Figure 3.29a, the number of transmissions is four compared to two in the uncompensated TWR case. Therefore, the wireless medium is occupied longer and not available for data transmission. In order to solve this issue, an approach shown in Figure 3.29b is usually used. In this approach, three transmissions are performed. Nevertheless, TWR is again performed

twice. Once using the first two transmissions, and for the second time using the second and the third transmission. The second transmission is actually used as a reply to node N_1 from node N_2 and also for initiating the second TWR.

Additionally, a slightly different Double Two Way Ranging (D-TWR) method is proposed in [66]. Using this method more robust ToA estimation can be achieved by further reducing the ranging errors due to crystal clock frequency offset.

An approach referred to as Symmetric N-Way Ranging (Symmetric NWR) is proposed in [101] to compensated the ranging error in an N-Way ranging approach. This method doubles the number of transmissions compared to N-Way ranging and compensates only the ranging errors between the neighboring nodes. The ranging errors between the non-neighboring nodes are left uncompensated.

In this work, in Chapter 5, a new approach for compensation of this error is proposed. It compensates the error, while at the same time keeps the number of transmissions at minimum.

Chapter 4

Modified equivalent time sampling - A method for reducing the range binning due to finite sample rate

The range binning effect can be a significant problem in ranging and localization applications, as shown in Section 3.4.3. The obvious solution is to use an oversampling A/D converter which, in some applications, does not justify the additional costs. In most wireless data transceivers, the used A/D converters are already oversampling the received waveform [12]. Nevertheless, the oversampling is usually only by a factor of two to four, which is not improving the ranging precision significantly. The methods discussed previously in Section 3.4.3 can reduce the ranging error due to finite sample rate, but require significant computational complexity. This limits their use for indoor localization where the used devices should have minimal complexity.

In this work we propose the **Modified Equivalent Time Sampling (METS)** method. It creates an oversampled version of the waveform used for ranging, without using oversampling data converters. It is derived from the known equivalent time sampling (ETS) method [17, 42, 72]. This method is intended to be implemented on standard digital wireless data transceivers using the available data converters and without additional hardware complexity added to the system.

4.1 Equivalent time sampling

Equivalent time sampling is a method for increasing the sample rate of a digitizing system with minimal increase of the hardware complexity. This approach is mainly used in high sample

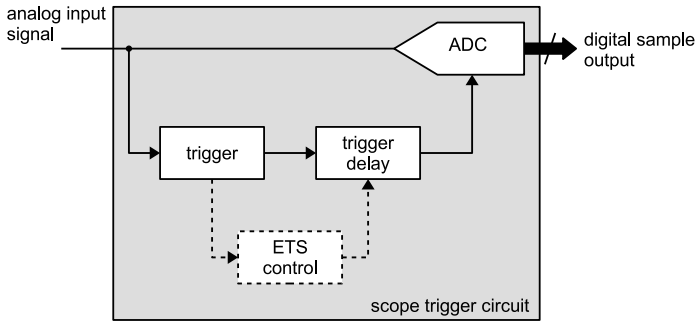


Fig. 4.1 Digital sampling oscilloscope triggering circuit and A/D converter

rate digital storage oscilloscopes (DSO). It works only for periodic input signals. The hardware complexity is basically in the triggering circuit of the oscilloscope. This method does not work and is not intended for sampling of aperiodic signals.

The basic idea behind the ETS is to sample each new period of a periodic signal in different time instants, relative to the previous period. In Figure 4.1, a simplified block diagram of a sampling circuit of a DSO is given. A triggering circuit (*trigger* in Figure 4.1) generates a pulse when a specific structure in the input waveform is detected. The trigger is used to detect, for example, the start of a new period of the periodic signal. After the trigger circuit, there is a delay block (*trigger delay* in Figure 4.1). It delays the trigger pulse for a specified amount of time. This extra delay is used when the signal after the *trigger delay* interval is of interest. This feature is available also on analog oscilloscopes.

The equivalent time sampling DSO has one more additional block: the *ETS control* block. This block actually controls the delay function automatically. In each new period of the periodic input signal, the delay is increased for a subsample interval. Therefore, instead of sampling each period in the same time instants, they are slightly shifted.

The ETS is further explained in Figure 4.2. As shown, in the first period, the start of sampling is not delayed. In the second period it is delayed for one subsample interval and for the next periods it is additionally increased. The final signal is reconstructed by interleaving the samples from the four periods, as shown in Figure 4.5. The number of periods used for sample acquisition is equal to the oversampling factor. Further decrease of the subsample delay used in the trigger circuit allows increasing of the effective sample rate.

This is an efficient solution for oversampling a periodic signal, mainly due to the simplicity and the cost of the system. The main disadvantage is that it works only for periodic signals. Also, the circuits used for generating the subsample delays should have small jitter, since it would increase the noise added to the observed signal.

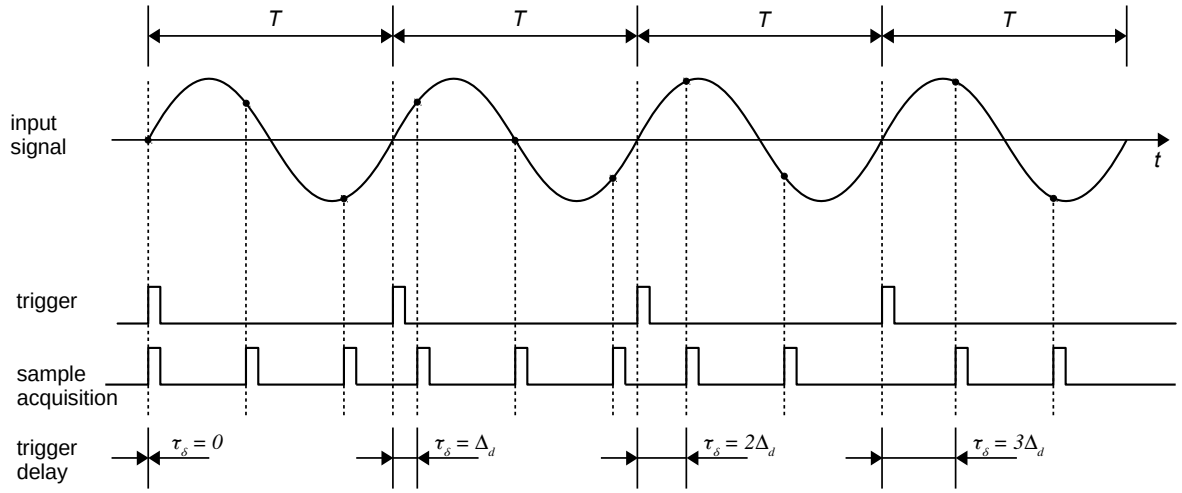


Fig. 4.2 Equivalent Time Sampling time diagram. In each period, after trigger conditions are met, the start of sampling is delayed for a given subsample delay.

Despite the few disadvantages this method is often implemented in many commercial products. The main motivation is the low cost, as well as the fact that in many cases periodic signals are of interest.

4.2 Modified equivalent time sampling (METS)

The ETS method, as previously described, cannot be used in ranging and localization applications which should be implemented on existing wireless data transceivers. The main reason is the inability to introduce subsample delays using the A/D converters present in the digital wireless receivers. For data transmission applications, subsample delays are not necessary at all and, therefore, direct implementation of ETS in the existing transceivers is not possible.

In this work, we propose an alternative approach for implementation of the ETS method, which we refer to as Modified Equivalent Time Sampling (METS). Unlike conventional ETS, the signal at the transmitter is prepared, in such a way, that no subsample delays at the receiver's A/D converter are needed. The goal is to shift the complexity towards the transmitter, i.e. to insert the subsample delays in the transmitted waveform. This approach would, at the end, lead to the same effect.

It should be mentioned that introducing subsample delays at the transmitter is not straightforward. The D/A data converters, used in digital wireless data transceivers, also do not have the capability of introducing a subsample delays. The implementation details of this method are given in Section 4.3.1. The waveforms that should be created at the transmitter and the pro-

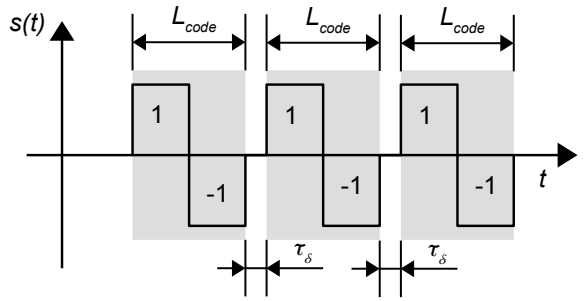


Fig. 4.3 Adding subsample delays between the three copies of a 2-bit PN sequence

cessing at the received required for reconstruction of an oversampled waveform are discussed here.

Structure of the transmitted waveform

To perform the ETS method, a periodic signal is needed. A periodic signal can be created by sending copies of the same waveform. Nevertheless, due to the lack of A/D converters with subsample delay capabilities, at the receiver, all of the received waveforms copies would be sampled at the same time instances. In order to solve this issue, we propose that the transmitter introduces subsample delays between each successive waveform copies being transmitted. With this modification, each copy of the received waveform, would be sampled at different time instances. An oversampled version of the original waveform can be reconstructed combining the samples from the different copies.

An illustration of this approach is given in Figure 4.3. In this example, the waveform to be transmitted consists of only two symbols[1] (in this case also bits). The first symbol is "1" (representing binary 1) and the second symbol is "-1" (representing binary "0"). For the sake of simplicity, only three copies are sent successively. Between the successive copies delays are inserted. These delays have a subsample duration of τ_δ. This signal, as shown in Figure 4.3, is not suitable for transmission, simply due to its infinite bandwidth. Therefore, the signal is filtered with a pulse shaping filter [64], usually a square root raised cosine filter. In Figure 4.4 an approximation of the filtered signal is shown. In order to transmit this signal, its samples are sent to the D/A converter. These samples, labeled with 1 to 7, are converted to analog samples using the D/A converter. The same D/A converter also performs an interpolation between the samples. With this interpolation, the same signal shown in Figure 4.4 is reconstructed,

[1] In practical implementations, the waveforms are slightly longer. Usually PN sequences of length 63 or more are used.

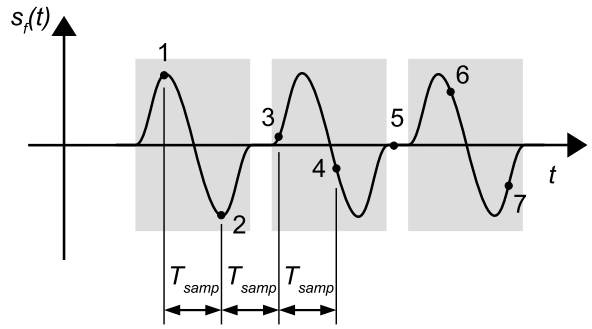


Fig. 4.4 Filtered version of the signal consisted of PN sequence copies with subsample delays between them

but in analog domain. This signal is further up-converted and transmitted, using a standard transmitter.

Reconstruction of the oversampled waveform at the receiver

The transmitted signal is received by the other wireless node, down-converted and sampled. The sampling is performed with a fixed sample rate of $f_{samp} = 1/T_{samp}$, where T_{samp} is the symbol interval length. For the sake of simplicity, it is assumed that the received waveform is sampled at the same time instants, labeled with 1 to 7 as in Figure 4.4. This is possible only if the data converters in both nodes are synchronized, which in real systems is not the case. Assuming perfect synchronization in this case, would not contribute to loss of generality and would make the explanation of the method much easier.

Even though the performed sampling is equidistant, the three received waveform copies are not sampled at the same time instants, due to the introduced subsample time delay τ_δ. In each successive waveform copy the sampling starts τ_δ seconds earlier, compared to the previous one. Since the waveform copies are sampled in different time instants, the acquired samples can be combined and interleaved, creating an oversampled representation of the waveform. This process is shown in Figure 4.5 where the samples from the three copies are interleaved in order to create one oversampled version of a single received waveform copy.

4.2.1 Model of the modified equivalent time sampling

To prove that the waveforms oversampled using the METS approach are the same as those obtained using oversampling A/D converters, models of the waveforms used at the transmitter and the receiver are derived.

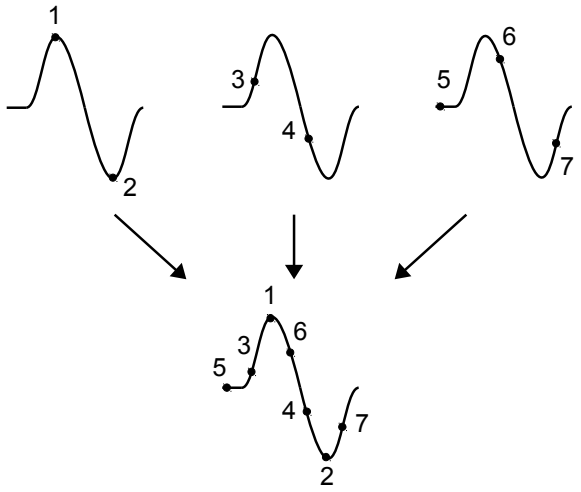


Fig. 4.5 Reconstruction of the oversampled waveform by interleaving the acquired samples

Transmitted signal model

The waveform to be transmitted is a digital waveform (e.g. PN sequence) consists of bits (chips) having values one and zero. These bits are mapped to symbols of a specific modulation (usually BPSK in ranging applications) and filtered using a pulse shaping filter. Therefore, for each bit a single pulse is obtained using the pulse shaping filter. The waveform obtained, consisting of an array of pulses, is represented as

$$s_{t1}(t) = \sum_{l=1}^{L} a_l h_f(t - lT_{sym}) \tag{4.1}$$

where L is the length of the waveform, a_l are the modulation symbols, $\{1;-1\}$ if BPSK is used, $h_f(t)$ is the impulse response of the pulse shaping filter and T_{sym} is the symbol interval.

The transmitter generates several copies which are shifted for an interval of τ_δ. The second waveform copy would be

$$s_{t2}(t) = \sum_{l=1}^{L} a_l h_f(t - lT_{sym} - \tau_\delta - LT_{sym}) \tag{4.2}$$

where τ_δ is the subsample delay introduced between the two sequences and the term LT_{sym} is present because the second waveform is delayed for L symbol intervals with respect to the first.

The m-th waveforms generated at the transmitter is given by

$$s_{tm}(t) = \sum_{L=1}^{L} a_l h_f(t - lT_{sym} - (m-1)\tau_\delta - (m-1)LT_{sym}) \tag{4.3}$$

The complete waveform containing M generated copies would be

$$\begin{aligned} s_t(t) &= \sum_{m=1}^{M}\sum_{l=1}^{L} a_i h_f(t - lT_{sym} - (m-1)\tau_\delta - (m-1)LT_{sym}) \\ &= \sum_{m=1}^{M}\sum_{l=1}^{L} a_l h_f(t - lT_{sym} - (m-1)(\tau_\delta + LT_{sym})) \end{aligned} \tag{4.4}$$

The waveform in (4.4) is up-converted and transmitted by the transmitter.

Received signal model

The transmitted waveform is received at the receiver. For a simplicity sake it is assumed that the channel is perfect and that the received waveform is the same as the transmitted one. The pulse shaping filter is usually split between the transmitter and the receiver, but we assume that the overall filter impulse response is contained in $h_f(t)$. The samples acquired from the waveform (4.1), for the first digital waveform copy would be

$$\begin{aligned} s_{r1}[n] = s_{r1}(nT_{sym}) = s_{t1}(nT_{sym}) &= \sum_{l=1}^{L} a_i h_f(nT_{sym} - lT_{sym}) \\ &= \sum_{l=1}^{L} a_l h_f((n-l)T_{sym}) \end{aligned} \tag{4.5}$$

For the second waveform (4.2) the acquired samples would be

$$\begin{aligned} s_{r2}[n] = s_{r2}(nT_{sym}) = s_{t2}(nT_{sym}) &= \sum_{l=1}^{L} a_l h_f(nT_{sym} - lT_{sym} - LT_{sym} - \tau_\delta) \\ &= \sum_{l=1}^{L} a_l h_f((n-l-L)T_{sym} - \tau_\delta) \end{aligned} \tag{4.6}$$

Finally, for the m-th waveform copy the acquired samples are

$$s_{rm}[n] = s_{rm}(nT_{sym}) = = \sum_{l=1}^{L} a_l h_f(nT_{sym} - lT_{sym} - (m-1)LT_{sym} - (m-1)\tau_\delta)$$
$$= \sum_{l=1}^{L} a_l h_f((n-l-(m-1)L)T_{sym} - (m-1)\tau_\delta) \tag{4.7}$$

The receiver interleaves the acquired samples to create the oversampled version of a single waveform copy. Therefore, the acquired sampled from the received waveforms copies ($m = 1 \ldots M$) should be first aligned. The alignment is performed by shifting samples from the m-th waveform copy by $(m-1)LT_s$ to left. The samples for the first waveform would not be shifted and would be

$$s_{rs1}[n] = \sum_{l=1}^{L} a_l h_f((n-l)T_{sym}) \tag{4.8}$$

The samples from the second waveform copy, shifted for LT_s would be

$$s_{rs2}[n] = s_{rs2}(nT_{sym} + LT_{sym}) = = \sum_{l=1}^{L} a_l h_f(nT_{sym} - lT_{sym} - LT_{sym} + LT_{sym} - \tau_\delta)$$
$$= \sum_{l=1}^{L} a_l h_f((n-l)T_{sym} - \tau_\delta) \tag{4.9}$$

For the m-th waveform copy the shifted and aligned samples would be

$$s_{rsm}[n] = s_{rm}(nT_s + (m-1)LT_{sym}) =$$
$$= \sum_{l=1}^{L} a_l h_f(nT_{sym} - lT_{sym} - (m-1)NT_{sym} + (m-1)NT_{sym} - (m-1)\tau_\delta)$$
$$= \sum_{l=1}^{L} a_l h_f((n-l)T_{sym} - (m-1)\tau_\delta) \tag{4.10}$$

Assuming that the $\tau_\delta = T_{sym}/N$, the waveform in (4.10) can be written as

$$s_{rsm}[n] = \sum_{l=1}^{L} a_l h_f((n-l)T_{sym} - (m-1)\frac{T_{sym}}{N})$$
$$= \sum_{l=1}^{L} a_l h_f((n-l-\frac{m-1}{N})T_{sym})$$
$$= \sum_{l=1}^{L} a_l h_f((\frac{nN-lN-m-1}{N})T_{sym}) \tag{4.11}$$

where $m = 1 \ldots M$.

Using the samples from (4.11), the oversampled version of a single waveform copy can be reconstructed. The oversampled waveform copy, $s_{ro}[k]$, obtained by interleaving of the aligned samples is given by

$$s_{ro}[nN - m - 1] = \sum_{l=1}^{L} a_l h_f((\frac{nN - lN - m - 1}{N})T_{sym}) \tag{4.12}$$

If the substitution $k = nN - m - 1$ is used in (4.12) the oversampled waveform copy $s_{ro}[k]$ would be

$$s_{ro}[k] = \sum_{l=1}^{L} a_l h_f((\frac{k}{N} - l)T_{sym}) \tag{4.13}$$

The oversampled waveform (4.13) is the final waveform produced by the proposed METS approach.

On the other hand, if the waveform $s_{t1}(t)$ from (4.1) is oversampled using an oversampling A/D converter, having sampling interval of T_{sym}/N, the obtained samples would be

$$\begin{aligned} s_o[k] = s_{t1}(k * \frac{T_{sym}}{N}) &= \sum_{l=1}^{L} a_l h_f(k\frac{T_{sym}}{N} - lT_{sym}) \\ &= \sum_{l=1}^{L} a_l h_f((\frac{k}{N} - l)T_{sym}) \end{aligned} \tag{4.14}$$

From (4.12) and (4.14) it can be concluded that the oversampled waveform obtained using the METS approach is equal to the one obtained using oversampling A/D converters.

4.3 Modified equivalent time sampling implementation

The METS method, as described in the Section 4.2, requires sub-sample delays to reconstruct an oversampled version of the waveform used for ranging. The data converters at the transmitter and at the receiver, cannot generate the required subsample delays. In this section, an implementation for overcoming this problem is proposed.

Using the proposed solution, the subsample delays are generated at the transmitter, before the samples are sent to the D/A converter. With this solution data converters capable of equidistant sampling can be used for implementation of the METS method.

4.3.1 Generating the METS waveform at the transmitter

Figure 4.6 shows a block diagram of the proposed solution used to generate the METS waveform at the transmitter. The waveform, with the subsample delays included, is generated in the digital domain and all of the blocks before the D/A converter in Figure 4.6 are digital.

The METS approach is general and can be used for oversampling different types of waveforms. The same transmitter presented here can be used for other types of waveforms, if the blocks *P* and *Mp* are substituted by a waveform generator which generates the required waveform.

Phase coded waveforms are usually used to implement ranging functionality on an existing data communication transceiver which supports phase modulation for data transmission. Therefore, the transmitter described in this section also uses phase coded waveforms.

To generate the METS waveform a PN sequence is first generated using the block *P* in Figure 4.6. The obtained bits are mapped using the symbol mapper *Mp*. The length of the generated PN sequence is L and chip rate of the generated chips is R_{PN}. The chip (or bit interval) would, therefore, be $T_{PN} = 1/R_{PN}$. The chips from this sequence are further mapped to symbols to produce a phase coded waveform. We assume BPSK modulation, since it is usually used in ranging applications. Each **1** from the PN sequence is mapped to **1** and each **0** is mapped to **-1**. The obtained symbol rate would be equal to the chip rate, R_{PN}.

After the generated phase coded waveform a delay of $\tau_\delta = (1/N) \cdot T_{PN}$ should be inserted. Therefore, the symbol rate of the phase coded waveform is first increased to $N \cdot R_{PN}$, by upsampling. The upsampling is performed at the block *U*. This block inserts $N-1$ zeros between each two samples of the phase coded waveform.

At the end of the upsampled phase coded waveform, one sample with value **0** is inserted, before the start of the next phase coded waveform copy. The control block *CL* controls the generation of the PN sequence copies as well as the generation of the zero sample from block *Z*, at the end of each phase coded waveform. The block *C* inserts the zero samples between the upsampled phase coded waveforms. This signal is further filtered with a pulse shaping filter *F*. After filtering, the signal is decimated by a factor of *M*, down to rate R_{TX}. The decimator keeps every *M*-th sample from the incoming signal, and discards the rest. An anti-aliasing filter before the decimator is not needed in this case since the signal bandwidth was not increased by upsampling and zero sample insertion.

Finally, the digital samples are converted to an analog waveform using the D/A converter. The used converter performs interpolation between the digital samples in order to reconstruct the signal between them.

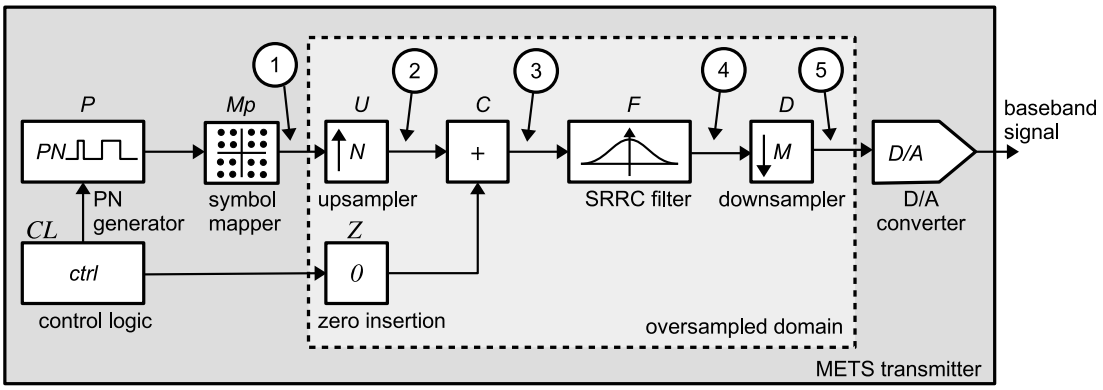


Fig. 4.6 Block diagram of the transmitter architecture used for introducing subsample delays between the successive copies of the PN sequences

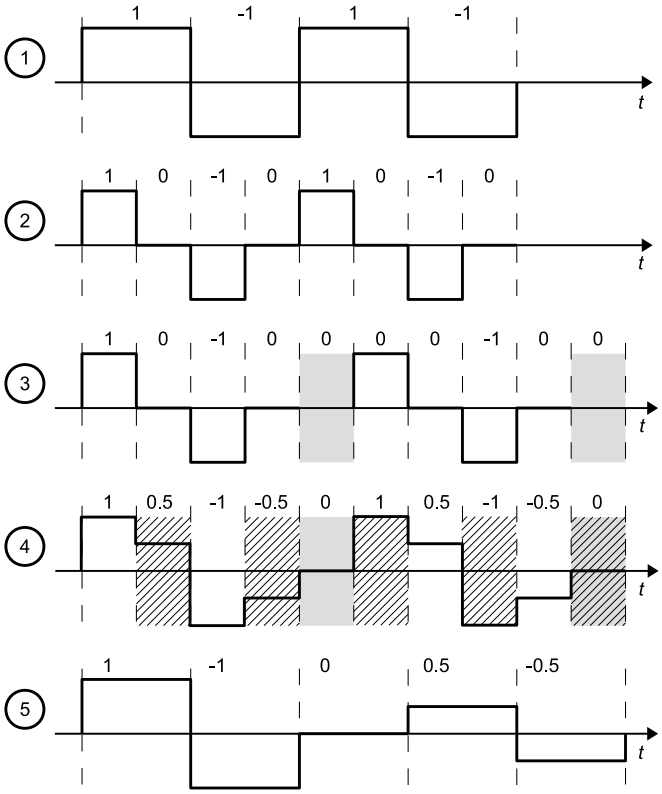


Fig. 4.7 Timing diagram of the waveforms in the METS transmitter

An example of METS generated waveform

To further clarify the functionality of the METS transmitter, shown in Figure 4.6, a simple example is presented here. The length of the used PN sequence, as well as the length of the filter are kept at minimum for simplicity sake. The same applies for the oversampling factor.

In this example, the used pulse shaping filter is a 2-tap filter. The impulse response of the filter is $h[n] = [1,\ 0.5]$. The PN sequence is $[1,\ -1]$. The oversampling ratio is $N = 2$. The downsampling ratio is $M = 2$. The output sample rate would be equal to the symbol rate ($R_{TX} = R_{PN}$). A total of two PN sequences are generated by the PN sequence generator "P".

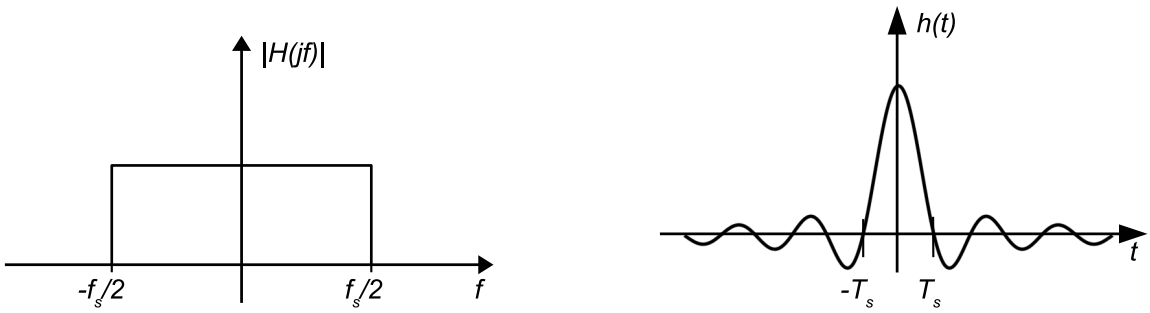
The numbers in the circles in Figure 4.6, are corresponding to the same numbers shown in Figure 4.7. They are used for referencing the different waveforms present in the block diagram. Figure 4.7 shows these waveforms. The vertical dashed lines in the same figure represent the chip, i.e. symbol intervals.

In Figure 4.7, the first waveform (1) is the output from the PN sequence generator, BPSK modulated using the mapper *Mp*. The two generated sequences are represented with a total of four chips, two for each sequence. One sample per chip of the sequence is used. The second waveform (2), is the output of the up-sampler. The number of sample intervals is doubled, due to the upsampling factor of $N = 2$. The up-sampler inserts one zero sample between the samples of the non oversampled signal. The third waveform (3) is the waveform obtained after adding the additional zero sample, at the end of each upsampled sequence. The added zero sample is highlighted in gray. This signal is further filtered with the pulse shaping filter, *F*. The output of the filter is the fourth (4) waveform. This waveform is further downsampled using the decimator *D*. The decimation factor in this case is $M = 2$. The decimation is performed, by keeping only the even numbered samples and discarding the rest. The discarded samples are shaded in the fourth (4) waveform in Figure 4.7. Finally, the signal labeled with 5 is obtained and sent to the D/A converter. This signal consists of 2 copies of the phase coded waveform. The second copy is time shifted for a subsample interval of $\tau_\delta = 1/2$.

This example slightly differs from a real scenario. The main difference is in the sample rate supplied to the D/A converter. The D/A converter generates one sample per symbol. This is usually not feasible in real implementations. The main limitation is imposed by the pulse shaping filter which requires at least two samples per symbol. Therefore, the ratio between the upsampling factor N and downsampling factor, M, should satisfy

$$\frac{N}{M} > 1 \tag{4.15}$$

In a real case, having a PN sequence of length L, upsampling factor of N and downsampling factor of M, would mean that a total of M sequences have to be generated. The total number of samples generated for one transmission by the PN sequence generator would be $L \times M$. After upsampling, the number of samples would be $N \times L \times M$. By adding a zero at the end of each sequence, a total of M zeros would be added to the complete METS waveform. Finally, after



(a) Frequency response **(b)** Impulse response

Fig. 4.8 Frequency and impulse response of an ideal low pass filter

filtering and downsampling by M, the number of samples that would be supplied to the D/A converter would be $(N \times L \times M + M)/M = N \times L + 1$. It is important to mention here that only integer ratios between N and M are investigated in this work.

The pulse shaping filter

The pulse shaping filter F in Figure 4.6 is commonly used in wireless data transceivers. It has two main roles: to limit the bandwidth of the transmitted waveform and to minimize the intersymbol interference caused by the channel. It can be also used as a matched filter, at the receiver, in order to maximize the signal to noise ratio of the received signal [110].

The METS transmitter described here also transmits digital data and must include pulse shaping filter for limiting the signal bandwidth, minimizing the intersymbol interference and maximizing the signal to noise ratio.

Different pulse shaping filters are available but the most popular are: the **rectangular** and the **raised cosine (RC)** pulse shaping filters.

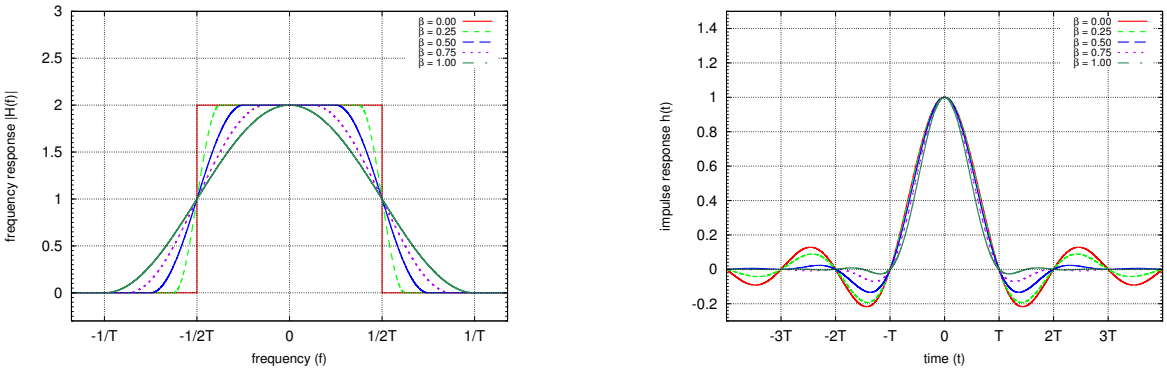
The simplest pulse shaping filter is the **rectangular** filter. The frequency and the impulse response of this filter are shown in Figure 4.8. In the frequency domain, the transition band[2] of this filter is infinitesimally small and, therefore, the guard intervals between neighboring channels in the frequency spectrum can be minimal.

Nevertheless, the tails[3] of the impulse response contain significant amount of energy, which can produce large PAPR at the transmitter.

The most commonly used pulse shaping filter is the **raised cosine (RC)** filter. It is also widely accepted in many wireless data transmission systems. The frequency response of this

[2]Each filter has a pass band, stop band and transition band. The transition band is the band between the pass and stop bands.

[3]Tails are called the ripples of the impulse response present around the main lobe (peak).



(a) Frequency response

(b) Impulse response

Fig. 4.9 Frequency and impulse response of a raised cosine filter.

filter, shown in Figure 4.9a is defined as

$$H_{RC}(f) = \begin{cases} T, & |f| \leq \frac{1-\beta}{2T} \\ \frac{T}{2}\left[1 + cos\left(\frac{\pi T}{\beta}\left[|f| - \frac{1-\beta}{2T}\right]\right)\right], & \frac{1-\beta}{2T} < |f| \leq \frac{1+\beta}{2T} \\ 0, & |f| > \frac{1+\beta}{2T} \end{cases} \tag{4.16}$$

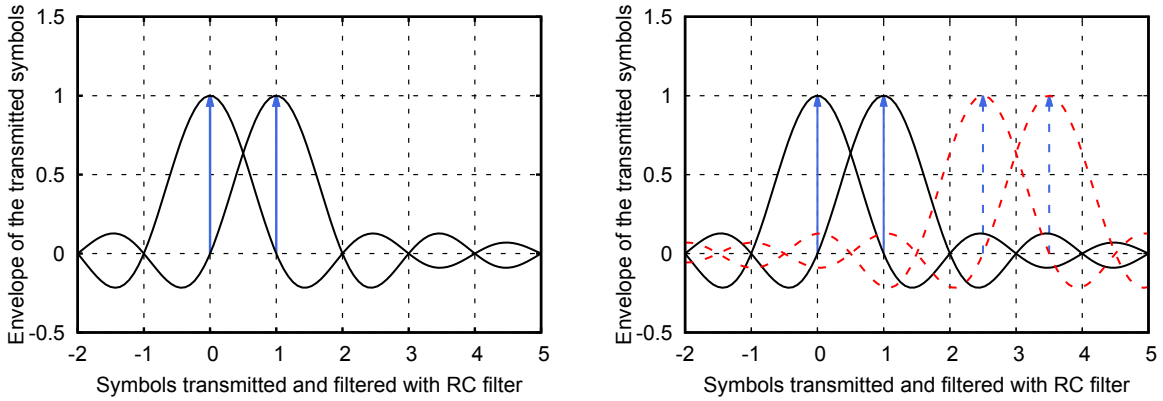
where $-1 \leq \beta \leq 1$ is the roll-off factor of the filter. The roll-off factor of the filter is also a measure for the excess bandwidth with respect to the Nyquist bandwidth. The impulse response of the filter, shown in Figure 4.9b, is given by

$$h_{RC}(f) = \begin{cases} \frac{\pi}{4} sinc(\frac{1}{2\beta}), & t = \pm\frac{T}{2\beta} \\ sinc(\frac{t}{T})\frac{cos(\frac{\pi\beta t}{T})}{1-(\frac{2\beta t}{T})^2}, & otherwise \end{cases} \tag{4.17}$$

The frequency response, as well as the impulse response of the raised cosine filter are given in Figure 4.9. In order to use the raised cosine filter as a matched filter, a square root version of it is used at the transmitter and at the receiver. This filter is called a square root raised cosine (SRRC) filter.

4.3.2 Inter-sequence interference

Due to transmission of a few copies of the same waveform, shifted by subsample intervals, an inter-symbol interference between the symbols of two successive copies would be present. This intersymbol interference is referred to as inter sequence interference, since the symbols from one sequence interfere with the symbols from the neighboring sequence. The reason for



(a) Symbols transmitted in regular intervals without subsample delays

(b) A delay of additional 0.5 samples added between the second and the third symbol

Fig. 4.10 Inter-sequence interference due to added subsample delay

the inter-sequence interference are the subsample intervals inserted between two successive sequences. The pulse shaping filter is not intended to minimize the intersymbol interference if the symbols are shifted for subsample intervals.

A single transmitted sequence, $[1,\ 1]$, filtered using rectangular pulse shaping filter is shown in Figure 4.10a. The arrows show the ideal sampling points, when the waveform is sampled with non-oversampling A/D converters. At the sampling points, the tails from the neighboring symbols have value of 0, i.e. no interference.

In Figure 4.10b, one more successive sequence, delayed for an additional half symbol interval, is shown with dashed lines. In this case, the first sequence is sampled again at the same points as previously. The tails coming from the symbols of the second copy, do not have value of 0 at the sampling points (arrows) of the first sequence. Therefore, each sample would contain parts from the tails of the symbols in the second copy.

Due to the inter sequence interference, using the METS as described previously, would not produce optimal results. The obtained samples would include additional inter-symbol interference (ISI) and would not represent the original signal.

To avoid the ISI coming from the successive copies, we propose having an additional guard interval (GI) between them. The guard interval length should be equal to the length of the impulse response of the used pulse shaping filter. Having this length of the guard interval, the tails from one sequence would not interfere with the neighboring sequence. The sub-sample delay should be added in this case after each GI. This guard interval would eliminate the ISI between the successive waveform copies and the oversampled waveform can be ideally reconstructed.

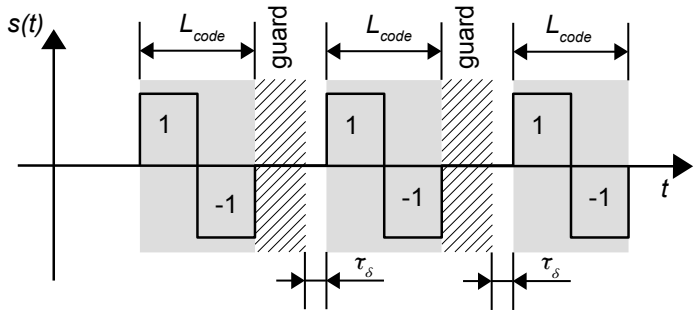


Fig. 4.11 METS waveform before filtering and transmitting. The hatched areas are the added guard intervals.

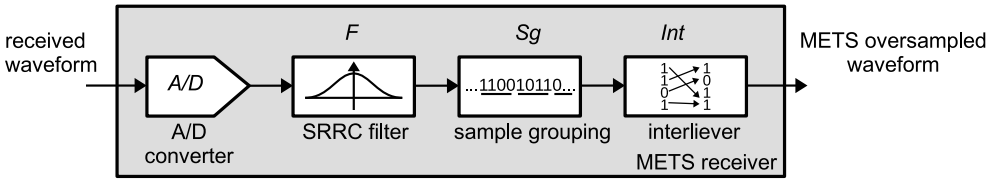


Fig. 4.12 Block diagram of a typical METS receiver

Figure 4.11 depicts the guard intervals inserted between the successive copies.

4.3.3 Reconstruction of the oversampled waveform at the receiver

Typically, the incoming METS waveform is sampled at the receiver with the sample rate[4] as the one used at the transmitter. The number of samples obtained at the receiver should be equal to the number of samples transmitted. These samples are used for reconstruction of the oversampled waveform. A block diagram of a typical METS received is shown in Figure 4.12

In order to keep the reconstruction process as simple as possible, it can be assumed that the guard interval, discussed in the Section 4.3.2, is a part of the phase coded waveform. Therefore, the length L of this waveform, would include the guard interval.

As explained in Section 4.3.1, the number of samples transmitted would be $N \times L + 1$. Ideally, it can be assumed that the samples obtained at the receiver are the same samples which were transmitted at the transmitter. This is very unlikely in a real scenario, but this problem would be further addressed in Section 4.3.4.

In order to reconstruct the oversampled signal, the received samples are first grouped. There is a total of M groups, where M is the downsampling factor in the transmitter. The first from the

[4]The sampling rate can differ between the transmitter and receiver.

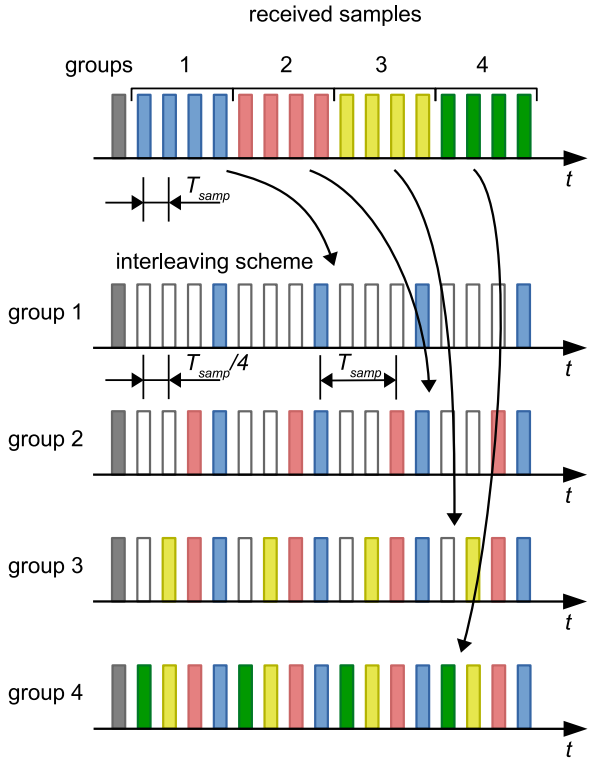


Fig. 4.13 Interleaving of samples at the METS receiver in order to reconstruct a four times oversampled waveform

received samples is used as the first sample of the reconstructed signal. The remaining samples are grouped in groups of $(N/M) \times L$ samples.

Further, the samples from different groups are interleaved. The process of interleaving is shown in Figure 4.13. In order to perform the interleaving, the samples of the first group are arranged in a specific way. In front of every sample from the first group, $N/M-1$ free spaces for new samples are reserved. Each sample from the second group is placed in front of the sample of the first group. The process continues and the samples from the next group are placed in front of the samples of the previous group.

An example of interleaving the received samples is shown in Figure 4.13. In this example the upsampling factor is $N = 4$, the downsampling factor is $M = 4$ and the sequence length is $L = 4$. The number of received samples is $N \times L + 1 = 17$. There are $M = 4$ groups labeled with numbers from 1 to 4. In Figure 4.13, the placement of the samples for each group is shown. The samples from the first groups are placed in such a way that there are three free spaces between the samples. These free spaces are drawn with white rectangles. The samples from the remaining groups are placed on the free spaces as previously explained. The interleaved

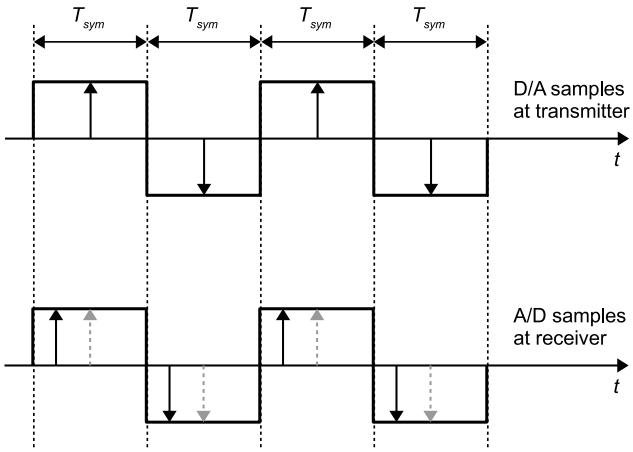


Fig. 4.14 Timing synchronization error

samples represent the oversampled version of the phase coded waveform, obtained using the proposed METS approach.

4.3.4 Synchronization of the transmitter and the receiver

In the proposed METS approach, for obtaining an oversampled version of the phase coded waveform, it was assumed that the data converters in the transmitter, as well as in the receiver, are perfectly synchronized. Nevertheless, in real scenarios, data converters in different wireless nodes are not synchronized.

The proposed approach can tolerate imperfect timing synchronization, but not imprecise frame synchronization.

Symbol timing synchronization

When the data converters in different nodes are not synchronized, it can occur that the A/D converter does not sample the received waveform in the center of the symbol interval. This issue is shown in Figure 4.14. The samples produced by the D/A converter at the transmitter are depicted with arrows. It can be assumed that these samples are placed at the center of the symbol intervals. In an ideal case, the samples produced by the A/D converter, at the receiver, would fall in the center of the symbol intervals, as shown with the dashed arrows. In a real case, the samples are not positioned in the center, as shown with the solid arrows.

The synchronization of the sampling with the symbol interval is called symbol timing recovery (or synchronization). In data communication systems the symbol timing recovery is

very important. Without correct symbol timing synchronization, decoding of the received data would not be optimal, leading to increased bit error rate.

This is not the case for the proposed METS approach. In the transmitted sequences, subsample delays are intentionally introduced. Therefore, if the symbol timing synchronization is well aligned for the first arriving sequence, it would not be aligned for the remaining sequences, since they are delayed for some subsample intervals. Therefore, aligning the samples to be obtained with the symbol intervals would not be necessary.

Frame synchronization

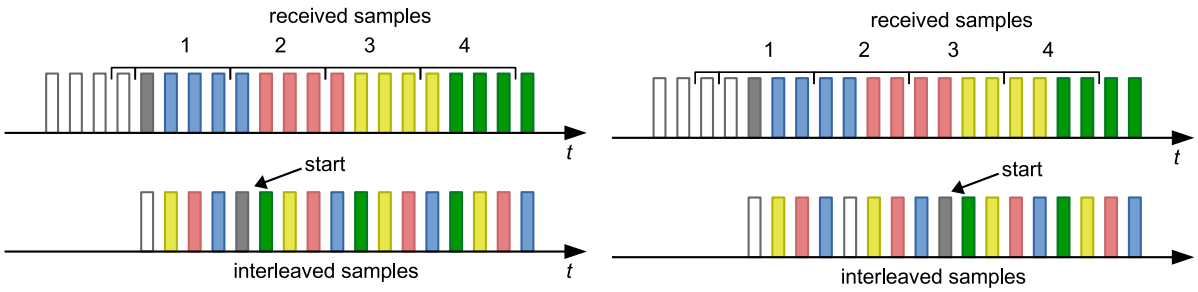
In the proposed approach the transmitted samples can be assumed to form a frame. The frame would consist of $N \times L + 1$ samples. The receiver should start sampling when the first transmitted sample arrives at the receiver. Although the METS approach is immune to symbol timing synchronization, it would not tolerate bad frame synchronization.

In order to perform frame synchronization, data receivers usually use an additional preamble in order to detect the start of an arriving frame. The implementation of a coarse synchronization for the frame start is relatively simple task, but it would not offer sample precision. It only serves for the purpose of detection of the incoming frame. When an incoming frame is detected, an additional fine synchronization is used in data transceivers to perform a sample precise synchronization. For ranging and localization purposes, usually only a coarse synchronizer is needed.

If the ranging and localization is integrated together with a data transmission system, time windows in which the wireless channel is assigned for ranging and localization purposes can be scheduled. These windows, should be synchronized between the wireless nodes involved in the process. The synchronization of these windows requires only coarse synchronization, and is easily achieved. The fine synchronizer is not needed, since the time of arrival estimation process is basically the same as a fine synchronization process. The difference is that the fine synchronizer should find the start of the frame with a sample precision and, on the other hand, the time of arrival estimator should offer a subsample precision.

In order to discuss the problem of synchronization we can assume that the samples, before and after the ranging frame, have a zero value. In a real case, there would be at least noise present. It can be also assumed, without loss of generality, that the sampling can only start earlier and not later than the start of the ranging frame. With a proper setting of the guard intervals, this can be easily achieved.

The receiver acquires $N \times L + 1$ samples. Due to coarse synchronization, the receiver can start the acquisition K sample intervals earlier. Therefore, the last K samples would not



(a) Acquisition started one sample earlier. **(b)** Acquisition started two samples earlier.

Fig. 4.15 Imperfect frame synchronization. Additional zero samples appear in the reconstructed waveform.

be acquired and included in the reconstruction of the oversampled version of the waveform. Instead, the leading zeros would be included in the acquired samples. They would later appear in the reconstructed signal.

The frame synchronization issue is shown in Figure 4.15. The squares in white color represent the zero samples before the ranging sequence. The samples in the other colors belong to the ranging sequence. As can be noticed they are grouped in groups of 4 samples plus the first sample. The oversampling factor N to be achieved is four.

In Figure 4.15a, the sampling starts one sample earlier. The acquired samples are interleaved and the waveform is reconstructed using the METS approach. It can be seen that the last sample is not included in the reconstructed waveform. Instead, there is an additional zero sample at the beginning of the reconstructed waveform.

In the second case, shown in Figure 4.15b, the acquisition starts two samples earlier. In the reconstructed (i.e. interpolated) waveform two zero samples instead of the two last samples, are present. If the acquisition starts even earlier, more zero samples are going to be included in the reconstructed waveform.

One interesting advantage of the proposed METS approach can be observed in Figure 4.15. It can be noticed that, when the acquisition starts one sample earlier, the start of the ranging sequence, marked with an arrow, is four samples delayed in the reconstructed sequence. For acquisition that starts two samples earlier, the start of the reconstructed signal is eight samples delayed. This shows that the proposed method is invariant to time shifts. If the acquisition starts earlier, the oversampled sequence would be delayed for the same number of samples, multiplied by the oversampling factor and, therefore, compensating for the earlier start.

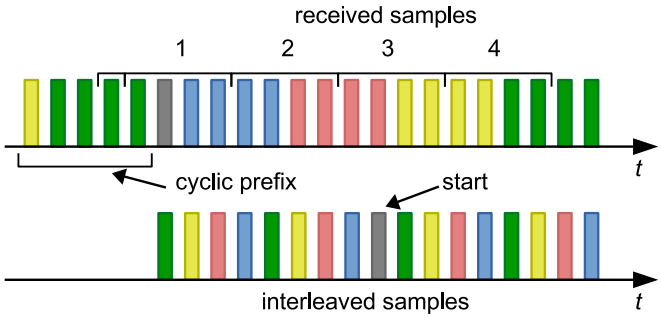


Fig. 4.16 Adding cyclic prefix for correcting synchronization errors. No zero samples appear in the reconstructed waveform.

Adding cyclic prefix for correcting coarse synchronization errors

In order to correct the coarse synchronization issue, two approaches are possible. The most obvious one is to use a fine synchronizer. As mentioned earlier, its implementation is complex and, therefore, avoided.

In this work, a solution using a cyclic prefix is proposed. The cyclic prefix must consist of $K < (N \times L + 1)$ samples, which are replicated from the end to the beginning of the original METS waveform. The cyclic prefix length must be smaller than the length of the original METS waveform in order to avoid ambiguity in the start of the reconstructed waveform.

An example of a METS waveform with added cyclic prefix is shown in Figure 4.16. In this example, the last five samples are copied to the beginning of the METS waveform, to form the cyclic prefix. The obtained sequence is transmitted. The receiver starts sampling two samples earlier, due to coarse synchronization with the frame start. The gray sample shown in Figure 4.16 should be the first sample acquired in an ideal case. The oversampled signal is reconstructed by interleaving the acquired samples. It can be seen in Figure 4.16 that the reconstructed signal does not contain zero samples. It can be also noticed that starting the sampling earlier is time invariant. The start in the oversampled sequence is delayed for eight samples, compensating for the earlier start. The eight samples of the oversampled waveform compensate exactly the two samples of the non-oversampled waveform.

In order to obtain the time of arrival, a cyclic correlation between the oversampled waveform and a locally generated oversampled copy of the same waveform, can be further performed.

4.3.5 Dynamic selection of the oversampling factor

When high oversampling factors are used, a large number of samples from the received waveform are acquired. Significant computational resources would be required for processing

of the acquired samples. These resources include memory for storing the samples as well as computational power needed for processing them. In the best case, the computational complexity would be increased linearly with respect to the number of samples.

For example, if pulse compression techniques are used, cross-correlation of the received waveform with a locally generated copy of the same waveform should be performed. Having an oversampling factor of N leads to a cross-correlation of N-times longer sequences. For low end devices like wireless sensors nodes this can be highly demanding. They would need large processing time and, therefore, a large amount of energy.

It is therefore important to have a mechanism for dynamic selection of the oversampling factor. The obvious approach is to negotiate the oversampling factor in advance. The nodes can agree on the oversampling factor before the ranging/positioning takes place. This solution is natural, but it has a few drawbacks. First, the negotiation must be performed, which requires extra frames to be transmitted on the wireless medium. Further, in a multi-user scenario, the negotiation of the oversampling factor should be performed between each pair of nodes. In a dynamic scenarios, where the SNR varies, the oversampling factor should be changed before each ranging/positioning takes place. This would introduce an enormous amount of transmissions needed for negotiation of the oversampling factor. Therefore, this approach is not optimal in multi-user scenarios.

In order to solve this issue, an extended version of the METS approach is proposed in this work. This extension is referred to as Binary Modified Equivalent Time Sampling (Binary METS).

Binary modified equivalent time sampling

As previously stated in Section 4.3.3, the acquired samples are grouped in M groups. The samples from one group are inserted before the samples from the previous group. Therefore, the samples from the next group are placed $\tau_\delta = T_{sym}/N$ seconds before the samples of the previous group, as shown is Figure 4.13. The time T_{sym} is the symbol interval length.

To optimize the selection process of the oversampling factor, a slight extension of the METS approach is proposed. The idea is that the samples from the next group are placed in the middle between the already reconstructed samples. With this change, the samples of the first group, as well as the first sample, are placed equidistantly at intervals of T_{sym}. The samples of the second group are placed in between the samples of the first group, on a distance of $T_{sym}/2$. Moreover, the samples of the third and fourth group are placed between the existing samples. The distance between the samples would be in this case $T_{sym}/4$. If more groups are available, the new samples are placed between the samples of the already processed groups.

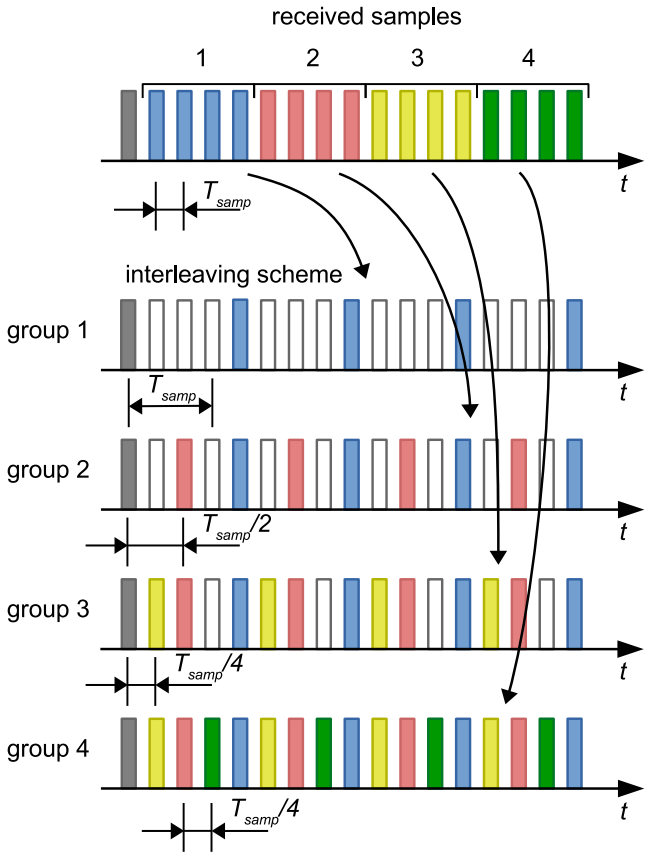


Fig. 4.17 Interleaving of samples using the binary METS approach in order to reconstruct a four times oversampled waveform

An example of the binary METS extension, for an oversampling factor of $N = 4$, is shown in Figure 4.17. The samples of the first group are placed at a distance of T_{sym}. The samples from the second group are placed in between the samples from the first group, at a distance of $T_{sym}/2$ from the already available samples. The samples from the third and fourth groups are placed between the samples from the first and second group, at a distance of $T_{sym}/4$.

The resulting waveform, obtained after the samples of the first group are processed, is not oversampled, i.e. one sample per symbol is available. With further processing of the samples from the second group, oversampling by a factor of two is achieved. Inserting the samples from the third and the fourth group, leads to an oversampling factor of four. Therefore, the receiver can choose how many groups are going to be processed (interleaved) and, therefore, dynamically select the oversampling factor. The approach is called ***binary***, because with each new step the oversampling factor is increased by two.

Generating the waveform for binary modified equivalent time sampling

To perform successive oversampling, using the binary METS approach, the transmitted samples must be generated in a slightly different fashion. The difference, compared to METS, resides in the insertion of zero samples in the oversampled domain.

If the maximal achievable oversampling rate is N, the number of the successive waveform copies, using the PN generator in Figure 4.6, would be again M. The samples from the second waveform copy, should be delayed for half a symbol interval, i.e. $T_{sym}/2$, with respect to the first generated waveform copy. In the oversampled domain this would correspond to insertion of $N/2$ samples between the first and the second upsampled waveform copy. The third and the fourth waveform copies, are used to reconstruct the samples which would fall between the samples from the first and the second group, when the signal is reconstructed at the receiver. Specifically, the samples from the third waveform copy fall between the samples of the first and second group. The samples from the fourth waveform copy fall between the samples from the second and the first group, as shown in Figure 4.17. Therefore, the samples from the third waveform copy should be advanced for $T_{sym}/4$ ($N/4$ samples in the oversampled domain) with respect to the second waveform copy, or delayed for $T_{sym}/2$ ($N/4$ samples in the oversampled domain) with respect to the first waveform copy. Advancing the third waveform copy with respect to the second waveform copy would produce overlapping between them. This can be fixed by inserting a guard interval of N samples between each two copies of the upsampled signal. The fourth waveform copy should be delayed for $T_{sym}/4$ ($N/4$ samples in the oversampled domain) with respect to the second waveform copy. This process continues by setting the delays in such a way that the samples from each new waveform would be placed between the already generated ones.

Synchronization of the binary modified equivalent time sampling

The binary METS approach also requires frame synchronization in order to reconstruct the oversampled waveform. Since the delay between the successive waveform copies is not constant, the previously proposed method with addition of a cyclic prefix, would not work.

A possible alternative is to increase the guard intervals between the successive waveform copies. The increase is performed by adding additional zero samples between the successive copies. The additional zero samples are also added before the first waveform copy.

In order to successfully reconstruct the oversampled waveform, the acquisition should start somewhere in the leading zeros, added before the first waveform copy. The number of samples in the formed groups should also be increased for the number of added zeros. The samples

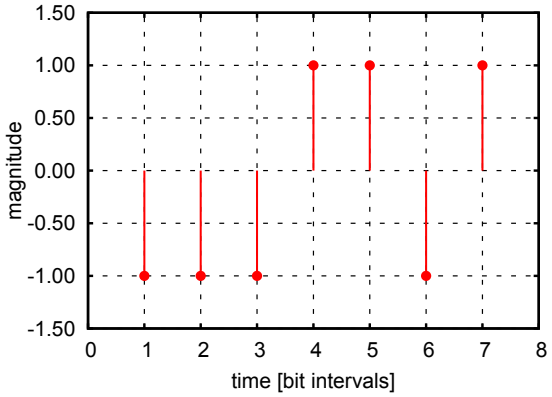


Fig. 4.18 Samples obtained by mapping m-sequence chips to BPSK modulation symbols

from these groups are further interleaved in the same fashion as for the described binary METS approach.

The guard interval must not be extremely large, since the coarse synchronizer can usually find the start of the frame within a few tens of samples. However, this approach adds more overhead, compared to the approach with the cyclic prefix.

4.4 Simulation of the modified equivalent time sampling

This section is devoted to the performed simulations, as well as the obtained results, of the proposed METS approach. A system consisted of a METS transmitter, shown in Figure 4.6 and a METS receiver, shown in Figure 4.12 is simulated under different conditions. The obtained results are compared with standard approaches like oversampling and interpolation of the received waveform.

The simulations were performed in MATLAB. The used PN sequence is a m-sequence with a length of 7 and generating polynomial $x^2 + 1$. The chips of this sequence are mapped to BPSK symbols in order to form the waveform that should be oversampled using the METS approach. The obtained samples are shown in Figure 4.18.

To apply the pulse shaping filter shown in Figure 4.6, the signal must be slightly oversampled. A SRRC filter with roll-off factor of 1 is used for pulse shaping. The chosen roll-off factor concentrates the energy in the main peak of the impulse response and keeps the energy in the tails at minimum. This ensures that the ISI between two successive waveform copies is kept at minimum, and also keeps the PAPR value low. The span of the filter is truncated to 6 sample intervals. It is implemented as a FIR filter and its impulse response is shown in Figure

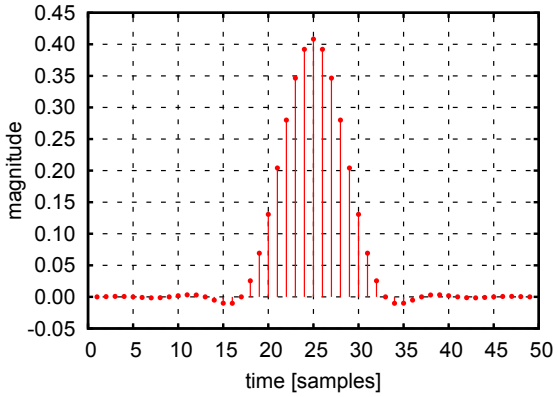
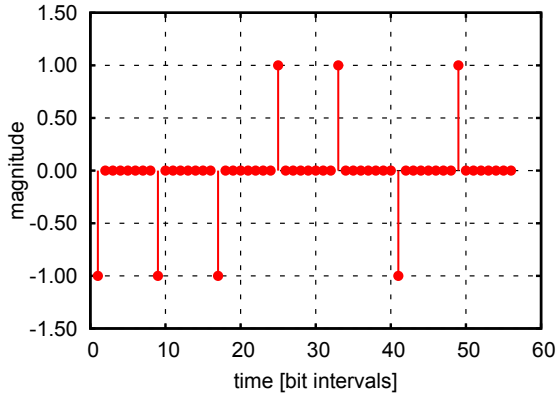


Fig. 4.19 Impulse response of a SRRC filter with $\beta = 1$, filter span of 6 samples and oversampling factor of 8.

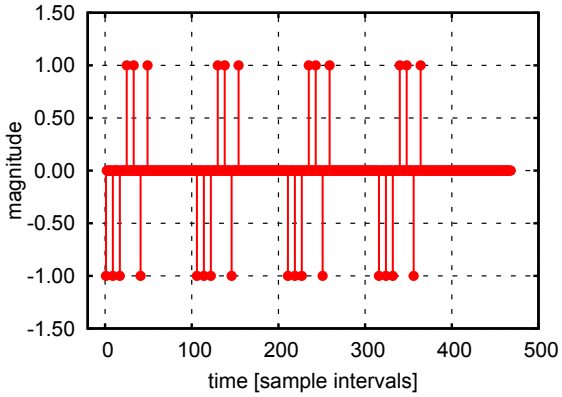
4.19. The oversampling rate of the filter is the same as the oversampling factor of the METS implementation.

The final oversampling factor, achieved with the METS approach in the performed simulation is $N = 8$. The upsampling (block U in Figure 4.6) is performed by inserting seven zero samples after each sample from the m-sequence. The total length of the upsampled sequence is $7 \cdot 8 = 56$ samples. The oversampled m-sequence is shown in Figure 4.20a. The sample rate at the output, after downsampling, would be 2 samples per symbol. The number of successive waveform copies would be $M = 4$. Between each two successive copies, a zero sample is added, in order to introduce the subsample delays. To avoid inter-sequence interference, a guard interval is introduced between the successive waveform copies. Since the RC filter span is 6 samples, the guard interval should be also $G = 6$ samples. These 6 samples are also upsampled by a factor of 8. The generated upsampled waveform, before filtering, is shown on Figure 4.20b. This signal is further filtered with the RC filter and downsampled. The filtered signal is shown in Figure 4.21b. A single oversampled and filtered m-sequence is shown in Figure 4.21a. The blue solid line on both figures is the envelope of the obtained waveforms. Downsampling is performed in the block "D" in Figure 4.6, by keeping every fourth sample coming out of the filter. The downsampled waveform is shown in Figure 4.22.

The downsampled waveform is transmitted through an ideal channel, to avoid both channel effects and noise, allowing a better evaluation of the approach. Anyway, in a relatively narrowband channel, it can be assumed that the frequency response of the channel is flat [7], under a given conditions. Therefore, the channel would be defined by an attenuation factor and noise, which does not affect this approach. If the channel is frequency selective, additional channel equalization can be performed, to minimize its effect.

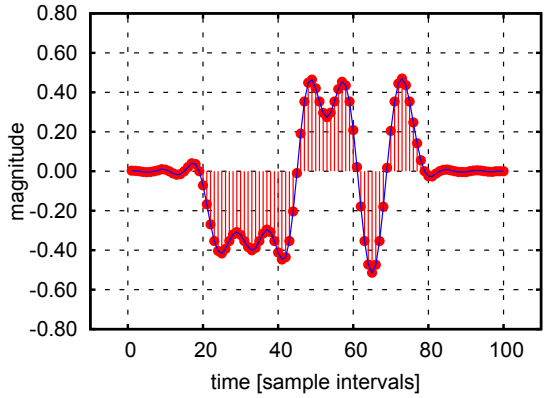


(a) *m*-sequence of length 7 up-sampled by a factor of 8

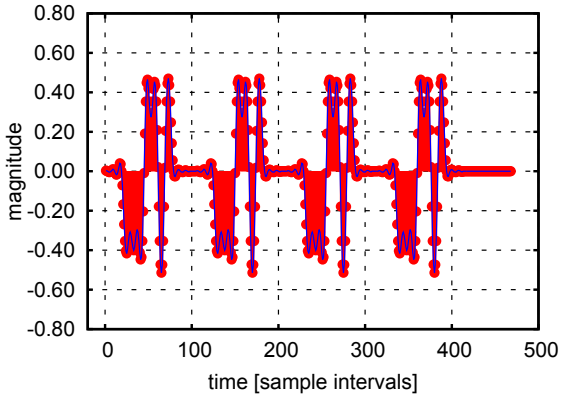


(b) Four copies of the up-sampled *m*-sequence with an additional guard interval between them

Fig. 4.20 Single and multiple up-sampled and non-filtered m-sequences



(a) Single filtered *m*-sequence



(b) Four filtered *m*-sequences creating a METS waveform

Fig. 4.21 A filtered version of the oversampled *m*-sequences

On the receiver side, the waveform is first filtered with a SRRC filter. The filter has the same roll-off factor as in the transmitter, but it is designed for a sample rate of two samples per symbol. The filtered samples are grouped and interleaved in order to perform the METS. In the simulation there would be $M = 4$ groups each containing $(L+G) \times N/M = (7+6) \times 2 = 26$ samples. There is one more additional sample in the complete METS waveform due to the M subsample delays, introduced at the METS transmitter. Finally, the reconstructed oversampled copy of the waveform is shown in Figure 4.23.

Figure 4.24 shows a comparison of the envelopes of the waveforms, oversampled using the METS approach and using an oversampling A/D converter. The two waveforms overlap and the small differences are due to the SRRC filter. The oversampled waveform is filtered

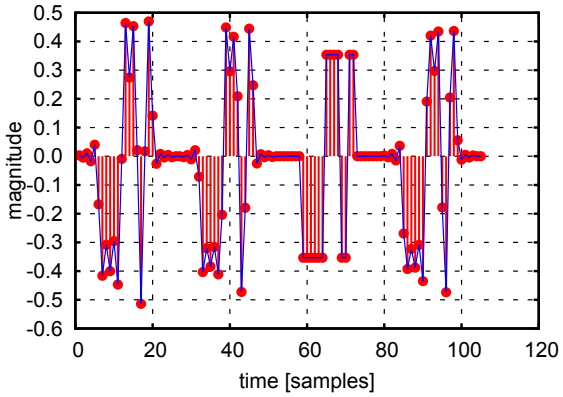


Fig. 4.22 Downsampled METS waveform ready for transmission

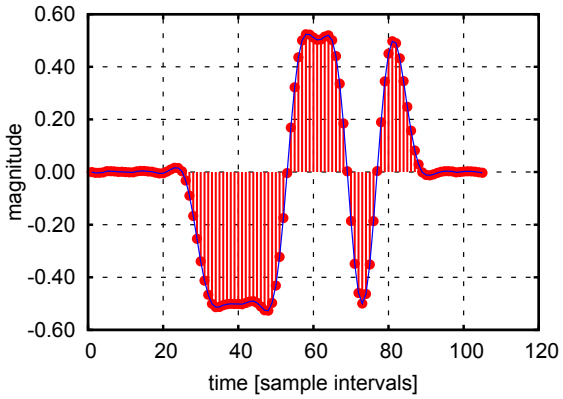


Fig. 4.23 Oversampled version of the waveform obtained at the receiver using the METS approach

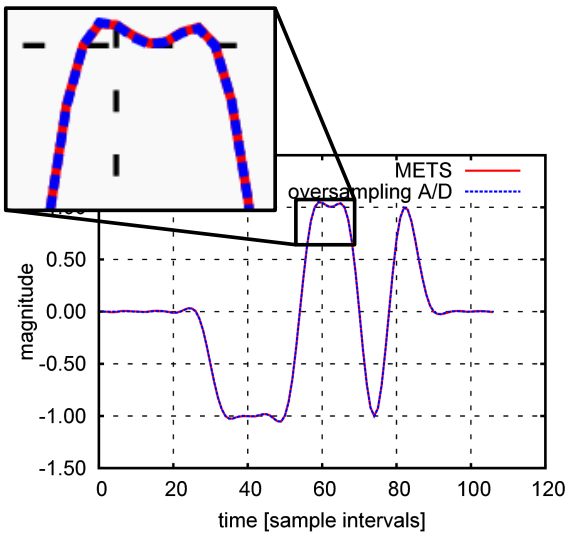


Fig. 4.24 METS waveform compared to the waveform obtained using oversampling A/D converters

with a SRRC filter designed for the oversampled domain (eight samples per symbol), while in the METS case the SRRC filter is for sample rate of two samples per symbol. The impulse response of the SRRC filter is more precisely represented, if the oversampling rate is higher, as when oversampling A/D converters are used.

When ranging or positioning is performed the oversampled version of the waveform, reconstructed using METS, is cross-correlated with the same waveform generated locally. The acquired peak in the correlation is used for estimation of the ToA, which is important in both the ToA based localization as well as in the TWR approach.

The simulations confirm the correctness of the proposed method. It was shown that in absence of noise and a frequency flat channel, the proposed method performs exactly as if oversampling data converters are used. More numerical results, as well as a comparison with other methods, are provided in Section 4.6.

4.5 Testing the modified equivalent time sampling on a software defined radio platform

The METS approach was tested and implemented in a real-life scenario using software defined radios (SDR). The generation of the METS waveform and the processing of the received samples was implemented in software. The METS waveform was transmitted from one radio to the other and the acquired samples were processed offline, since a real-time implementation would require huge effort and is out of the scope of this work.

The received waveform was also oversampled, using oversampling A/D converter, and compared to the waveform obtained using the METS approach.

4.5.1 The software defined radio platform

The software defined radio (SDR), used for testing of the METS approach, is an Ettus Research USRP N210 radio. This SDR supports sample rates of up to 50 MSps for the in-phase (I) and the quadrature (Q) signals. Only 8 bits per I and Q sample are possible for the highest sample rate and 16 bits for sample rates of up to 25 MSps. An analog front-end for the 2.4/5 GHz ISM band was used. The transmissions were performed in the 5 GHz band, since it is not so heavily used compared to the 2.4 GHz band. The two SDRs were separated at a distance of 1.5 to 2 meters. Both radios were equipped with omni-directional antennas having 3 dBi gain.

The transmitter used a sample rate of 5 MSps. Each symbol was oversampled by a factor of two. The receiver was sampling with a rate of 20 MSps. It was oversampling the

Table 4.1 Parameters of the SDR platform used for testing of the METS approach

Parameter/Specification	Value
SDR platform model	Ettus USRP N210
Maximum ADC sample rate	25 (50) MSps
ADC sample width	16 (8) bits per I/Q
Maximum DAC sample rate	25 (50) MSps
DAC sample width	16 (8) bits per I/Q
Analog front-end model	Ettus XCVR2450
Analog front-end bands	2400-2500 & 4900-6000 MHz
Maximum output power	100 mW
Carrier frequency	5750 MHz
Receiver sample rate	20 MSps
Transmitter sample rate	5 MSps
Transmit power	10 mW
Receive gain	30 dB
Antenna gain (omni)	3 dBi

received waveform by a factor of 8 samples per symbol. Additional decimation (filtering and downsampling) was performed in software in order reduce the sample rate to 5 MSps. The details for the used parameters in the performed test with the SDR platform are summarized in Table 4.1.

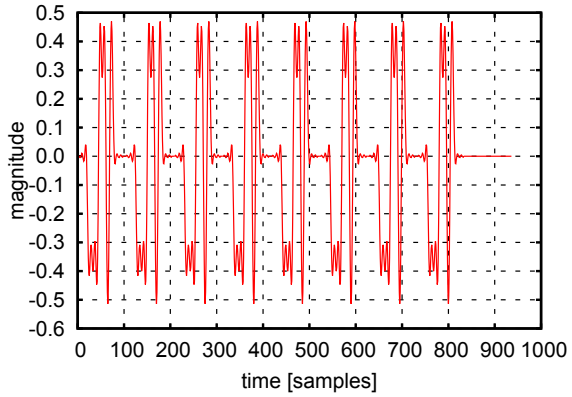
The transmit power and the receiver gain were set up in such a way that the acquired waveform was well above the noise floor and, at the same time, it had a significant dynamic range. The receiver acquired the data for one second and the transmitter was transmitting the same sequence repeatedly, each 100 milliseconds.

4.5.2 Generating the METS waveform

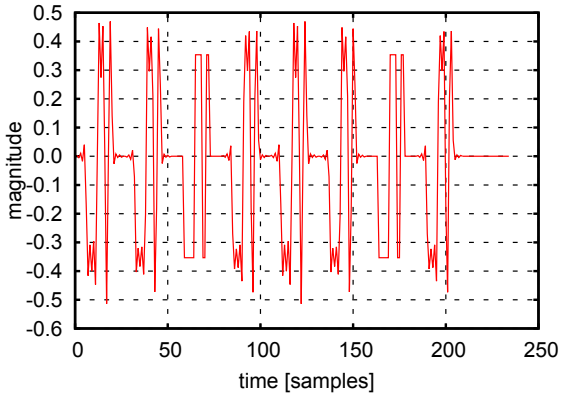
In order to prepare the signal for transmission with the SDR, an m-sequence of length 7 is first generated. The characteristic polynomial is $x^2 + 1$. A short sequence was used, in order to easily evaluate the proposed approach.

The generated m-sequence is first upsampled by a factor of $N = 8$. This is the final oversampling factor, which should be obtained after the waveform is reconstructed using the METS approach. The used modulation is BPSK, i.e. one symbol per bit.

A SRRC filter with a roll-off factor $\beta = 1$ is used. The filter length is 6 taps in the non-oversampled domain. A length of 48 taps is achieved with oversampling by a factor of 8.



(a) Four copies of the PN sequence with additional cyclic prefix of four copies, oversampled by 8 and filtered with SRRC filter.



(b) Downsampled version of the signal, prepared for transmission.

Fig. 4.25 Envelope of the oversampled and downsampled version of the METS signal.

The upsampled signal is further filtered with the SRRC and downsampled by a factor of $M = 4$. There would be no aliasing present, since the roll-off factor is selected with respect to the rate of the code, i.e. the Nyquist criteria is satisfied. The envelope of the generated waveform is shown in Figure 4.25. This waveform is further up-converted and transmitted using the SDR platform. Assuming that the baseband waveform, i.e. the METS waveform, is $\bar{s}(t)$, the up-converted waveform would be

$$m_{tx}(t) = \bar{s}(t) \cdot e^{j2\pi f_{0TX} t} \tag{4.18}$$

This is the so called phasor representation of the waveform. The waveform $\bar{s}(t)$ is a complex baseband signal representing the in-phase and quadrature components and e is the base of the natural logarithm. The frequency f_{0TX} is the carrier frequency, t is the time and j is the imaginary unit.

4.5.3 Processing of the received signal

After transmission, the waveform in Figure 4.25b is acquired at the receiver with a sample rate of N/M times the code rate of the PN sequence. This is the same sample rate at which the waveform was transmitted.

Before reconstruction of the oversampled waveform is performed, the frequency and phase offset of the received signal should be corrected. Without correction of the frequency offset the

reconstruction would not be possible. On the other hand, the phase offset correction can be performed after reconstruction of the oversampled waveform.

Frequency offset correction

The carrier frequency of the wireless nodes always differs from the nominal carrier frequency. This is due to the frequency inaccuracy of the clock sources used in the nodes. Therefore, the transmitter at one node would transmit at a carrier frequency slightly different compared to the carrier frequency of the receiver at the other node.

In order to demodulate the received signal, $m(t)$ in Equation 4.18, it is multiplied by a phasor $c_{RX}(t) = e^{-j2\pi f_{0RX}t}$. The frequency f_{0RX} is the carrier frequency at the receiver. The obtained signal would be

$$\begin{aligned} s_{rx}(t) = m(t) \cdot c_{RX}(t) = m(t) \cdot e^{-j2\pi f_{0RX}t} = \\ = \bar{s}(t) \cdot e^{j2\pi f_{0TX}t} e^{-j2\pi f_{0RX}t} = \\ = \bar{s}(t) \cdot e^{j2\pi (f_{0TX} - f_{0RX})t} = \\ = \bar{s}(t) \cdot e^{j2\pi \Delta f_0 t} \end{aligned} \tag{4.19}$$

Due to the frequency offset between f_{0TX} and f_{0RX}, the signal would not be completely demodulated. Modulation with the frequency offset $\Delta f_0 = f_{0TX} - f_{0RX}$, would be present. It would cause constant rotation of the constellation diagram of the received signal, in this case rotation of the received BPSK symbols.

The frequency offset should be compensated in order to minimize its effect on the METS approach. The compensation is performed by additional multiplication of the signal by $e^{-j2\pi\Delta f_0 t}$. Before frequency offset compensation is performed, Δf_0 must be estimated. Different frequency offset estimation algorithms exist [93, 100]. A maximum likelihood (ML) carrier frequency offset (CFO) estimator is used for frequency offset estimation [119] in this implementation.

Phase offset correction

After correcting the frequency offset, i.e. $\Delta f_0 = 0$ in (4.19), the phase offset is estimated and corrected. The phase offset is due to the different phase offsets of the carrier at the transmitter and at the receiver. It is not included in (4.19), in order to keep the equation simple. The demodulated signal, with frequency offset fully compensated and phase offset included would

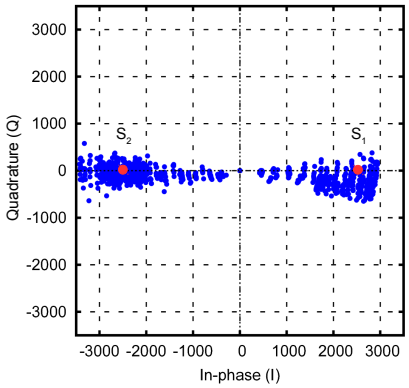


Fig. 4.26 Constellation plot of the complex samples of the received BPSK modulated signal

be

$$s_{rx}(t) = \bar{s}(t) \cdot e^{j(2\pi\Delta f_0 t + \varphi_0)} e^{-j2\pi\Delta f_0 t} = \\ = \bar{s}(t) \cdot e^{j\varphi_0} \tag{4.20}$$

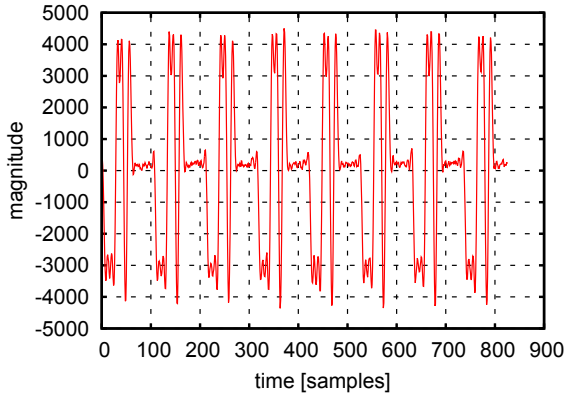
The constellation diagram, representing the received symbols, would be rotated for φ_0, which is the phase offset in (4.20). The phase offset correction would rotate back the constellation making $\varphi_0 = 0$.

In Figure 4.26 a constellation diagram of the signal received using the SDR, after carrier frequency and phase offset corrected, is shown. The modulation is BPSK. The samples are not grouped around two points representing the symbols $S0$ and $S1$ due to the oversampling of the received symbols. Some of the acquired samples are at the beginning of the symbol interval, some in the middle and some at the end. Those at the beginning and end of the symbol interval would have values closer to the center and those in the middle of the symbol interval would be further from the center of the constellation diagram.

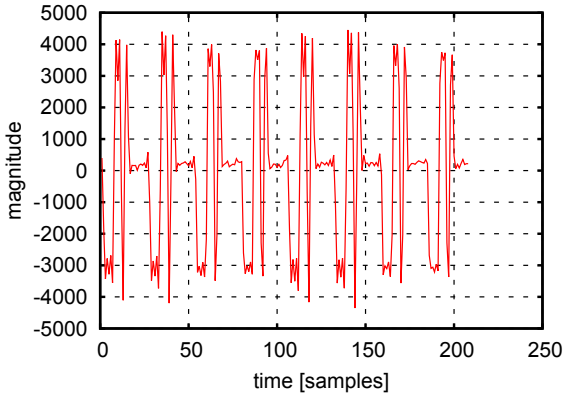
The quadrature component of the signal is not completely removed, as shown in Figure 4.26. The main reason is the noise, but also a small DC offset can be observed in the I and Q signals. They are relatively small and would not significantly affect the performed tests.

Reconstruction of the oversampled signal using the METS approach

The signal, obtained after frequency and phase offset correction, is shown in Figure 4.27a. This is only the in-phase component of the signal. The quadrature component of the signal is not taken into account. In an ideal case it should be equal to zero. It shows eight subsample shifted copies of the waveform obtained form the generated PN sequence. This waveform is already oversampled and is used for comparison with the waveform obtained using the METS

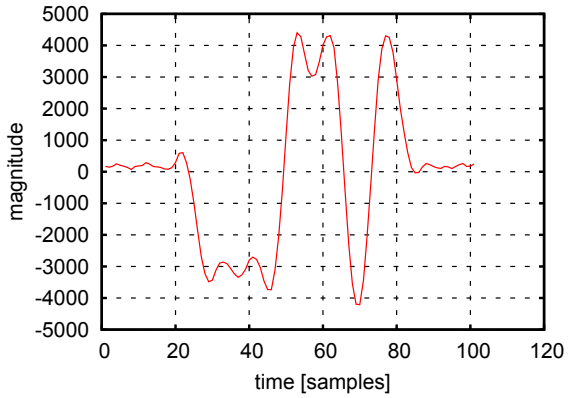


(a) Received BPSK signal oversampled by a factor of 8

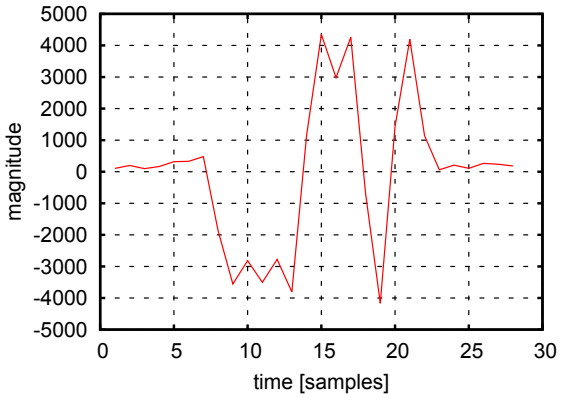


(b) Downsampled version of the received signal. The downsampling factor is four, leading to two samples per symbol.

Fig. 4.27 Oversampled and non oversampled version of the received signal



(a) Single copy sampled with eight samples per symbol.



(b) Single copy sampled with two samples per symbol.

Fig. 4.28 Single copy of the received signal consisted of m-sequences. Difference between the eight times oversampled sequence and sequence sampled with two samples per symbol is shown.

approach. The waveform shown in Figure 4.27b is the downsampled waveform, having a sample rate of two samples per symbol. This waveform is used for the METS approach. There are 8 instead of 4 waveform copies, due to the inserted cyclic prefix, used to avoid additional fine synchronization.

In Figure 4.28, only one waveform copy from the signals shown in Figure 4.27 is plotted. These signals are not yet filtered using the SRRC filter at the receiver and, therefore, more ripples can be observed. The waveform oversampled 8 times, has more details compared to the one used as an input signal for the METS.

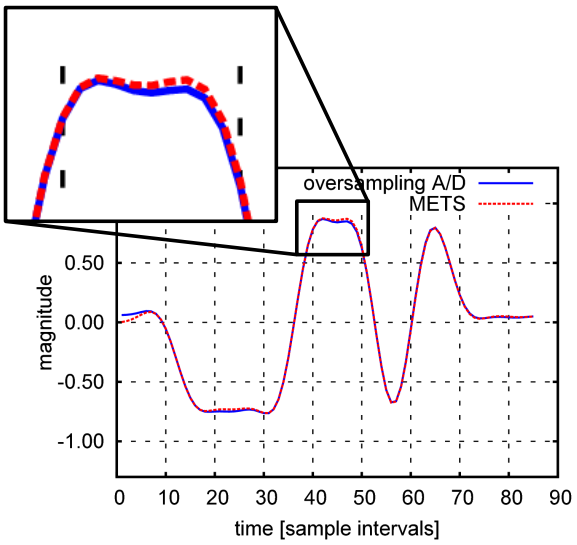


Fig. 4.29 Oversampled compared to the METS reconstructed waveform. The samples are acquired using the SDR platform.

The downsampled copy of the received waveform is further used to reconstruct an oversampled version using the METS approach. The reconstructed waveform is shown in Figure 4.29. It can be seen that there is an almost perfect match of both waveforms. The noise and the imperfections in the radios introduce slight differences between the two waveforms.

4.6 Performances of the METS approach compared with the oversampling and interpolation methods

In this section, a few different methods for obtaining oversampled version of the received waveform are compared with the proposed approach, in simulation. The METS approach is intended to be used in ranging methods and, therefore, the ranging precision obtained using this approach was evaluated. The range estimate obtained using a waveform oversampled by an oversampling A/D converter is used as a reference value. The ranging precision obtained using the METS approach is compared to the ranging precision obtained by different interpolation methods, i.e. : linear interpolation; nearest, next and previous neighbor interpolation; shape-preserving piece-wise cubic interpolation; spline interpolation and Fourier transform zero padding interpolation. These methods are selected since their computational complexity is comparable to that of the METS approach.

To perform the necessary simulations, a BPSK modulated waveform is created using the METS transmitter in Figure 4.6. The D/A converter, used to interpolate between the samples,

performs a $\frac{sin(x)}{x}$ interpolation with truncated impulse response. No noise is added. This waveform is assumed to be received at the receiver and samples using the oversampling A/D converter are obtained. The acquired samples are further downsampled to the same sample rate as the one used at the D/A converter. These samples are used for reconstruction of the oversampled waveform with the METS approach. In order to test the other interpolation methods, only the samples from the first waveform copy, obtained at the METS transmitter, are used. For all of the performed calculations a floating point arithmetic is used.

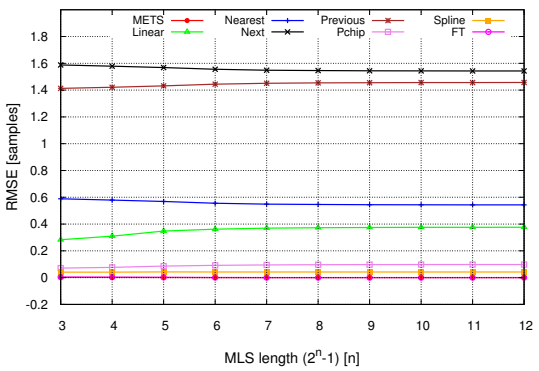
The obtained oversampled versions of the received waveform are used for ToA estimation. The ToA estimator is similar to the one shown in Figure 3.22. The first maximum search is performed using a ML estimator. The obtained ToA value is compared to the ToA reference value, obtained using the waveform acquired with the oversampling A/D converter.

Comparison of ToA root mean square error

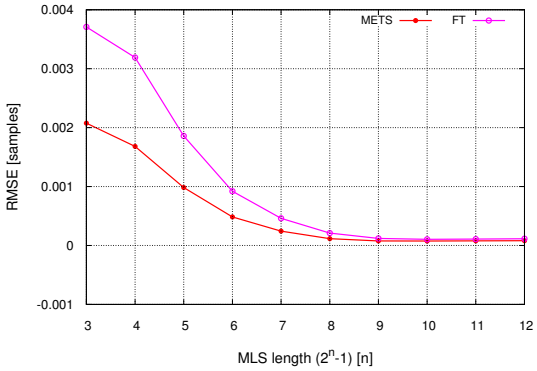
Multiple simulations were performed in order to obtain the ToA estimation RMSE values for the different methods. The same family of m-sequences (PN sequences) are used for all simulations. A total of 1000 simulation for each case are performed and the results are averaged. For each simulation run a different m-sequence, generated using the same generator polynomial, was used. Therefore, dependency of the obtained results from the used m-sequence is avoided. The results obtained using different interpolation methods, as well as, the proposed approach are compared to the reference values.

At first, the ToA RMSE was estimated as a function of the m-sequence length. The length is equal to $2^n - 1$ and n is changed from 3 to 12. The METS transmitter used in the performed simulation has an oversampling factor of $N = 8$ and a downsampling factor of $M = 4$. The filter is SRRC with roll-off factor of 1 and impulse response length of 6 symbol intervals. In Figure 4.30a, the plots for all the simulated methods are shown. The X axis is the m-sequence length, i.e. the exponent n. The Y axis is the RMSE of the ToA estimation in samples (in oversampled domain). The best results are achieved with the METS approach and the Fourier transform (FT) zero padding interpolation [75]. In Figure 4.30b, only the two methods with the lowest ToA estimation RMSE are shown. It can be noticed that the METS and the FT interpolation perform much better than the other interpolation methods. The METS approach is slightly better compared to the FT zero padding interpolation. According to the obtained results, the METS RMSE is half compared to the RMSE in the FT zero padding interpolation, i.e. it offers two times better precision.

Further, the dependency of the RMSE as a function of the SRRC filter length is evaluated. The METS approach is simulated for different lengths of the filter response. As shown in Figure



(a) RMSE of the ToA of all methods tested.



(b) RMSE of the ToA of the methods having lowest error.

Fig. 4.30 RMSE of the time of arrival estimation as a function of MLS length

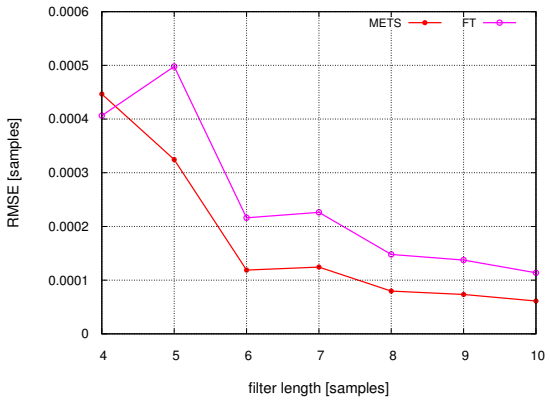


Fig. 4.31 RMSE of the ToA estimation as a function of SRRC filter length.

4.31, if a SRRC filter is used as a pulse shaping filter, the reduction of the ToA estimation error does not significantly improve with the length of the filter. The main reason is that for a roll-off factor of 1, the tails of the filter impulse response for lengths larger than 6, do not carry a significant amount of energy.

As shown in the results the FT zero padding interpolation achieves an error comparable to the METS approach. Nevertheless, if the FT zero padding interpolation is used in systems supporting fixed point arithmetic, worse results are expected.

Chapter 5

Crystal clock frequency offset compensation in N-Way ranging

In multi-user scenarios, localization estimation of more than a few wireless nodes is usually performed. ToA based methods are the best choice for these scenarios, since they introduce the smallest overhead needed for location estimation of a large number of nodes. In a ToA approach the same signal transmitted from the anchor nodes can be received and used by a practically unlimited number of nodes which require location estimation. The main disadvantage of ToA based methods is the high required synchronization precision of the anchor nodes. The clocks in the anchor nodes are inaccurate and, therefore, synchronization should be regularly performed to compensate the introduced clock drift. In order to avoid frequent synchronization of the anchor nodes, atomic clock generators can be deployed, e.g. GNSS [58]. These methods are mainly intended for use in infrastructure-based localization scenarios.

The TWR approach, on the other hand, has significantly lower synchronization requirements and can be used in infrastructure-based and infrastructure free, i.e. cooperative localization scenarios. In an infrastructure-based localization, a set of anchor nodes with known locations are used as reference nodes. They perform TWR with each node requiring a location estimate. The TWR between this node and the anchor nodes is performed in round robin fashion. A minimum of four anchor nodes are needed in order to estimate the position (x, y, z) of a node in a three dimensional space. Therefore, for each node, $2 \times 4 = 8$ (two per anchor node) transmissions are needed for a single location estimate. If more than four anchor nodes are available, the position estimate can be additionally improved by estimating the distances to the additional anchor nodes. This would additionally increase the number of required transmissions. Increasing the number of transmissions as well as increasing the number of nodes leads to high usage of the wireless medium for localization purposes. This can greatly restrict the

availability of the medium for data communication. If the localization infrastructure is intended to serve a large number of users, this can be a limiting factor in many applications, since it can significantly reduce the data transmission capacity.

Infrastructure free localization using TWR is performed by measuring the distances between all of the nodes in a wireless network. The obtained distances are used for relative position estimation of the nodes. In networks with more than a few nodes, performing TWR between all of the nodes can be challenging. The number of required transmissions is given as

$$N_{TX} = N(N-1) \approx N^2 \tag{5.1}$$

where N is the number of nodes in the network. In large wireless networks, the number of transmissions required for estimating the relative locations of the nodes can become significant. In order to solve this issue, the N-Way (or Multi Way) Ranging (NWR) is proposed in [40]. This solution requires N transmissions for localization of N nodes, at an expense of higher complexity. It is also sensitive to clock frequency offset due to the large reply time of the nodes. ToF estimation errors in the order of tens of nanoseconds are expected, for a typical clock frequency accuracy of $10-50$ ppm and millisecond reply interval of the nodes. These errors are comparable to the typical ToF values in indoor environments. In the N-Way Ranging (NWR) approach, these errors are additionally accumulated, making the method practically unusable, when ordinary crystal clock generators are used.

The ranging errors due to clock frequency offset, in N-Way ranging, are significantly larger compared to those in the TWR approach. The method is performed in such a way that it accumulates the ranging errors produced in each transmission and unusable ranging results can be produced.

In this work a new approach for compensation of ranging errors due to clock frequency offset in NWR method is proposed. The two main advantages of this approach are: first, the number of transmissions is kept at minimum, as in the basic NWR; and second, it uses the already available information of the frequency offset to correct the distance estimation errors. The information of the frequency offset is already available in wireless data transceivers. Therefore, the same frequency offset estimate can be used if the NWR is implemented in a wireless data transceiver. Nevertheless, if the NWR is implemented standalone, a separate estimator for frequency offset estimation must be included.

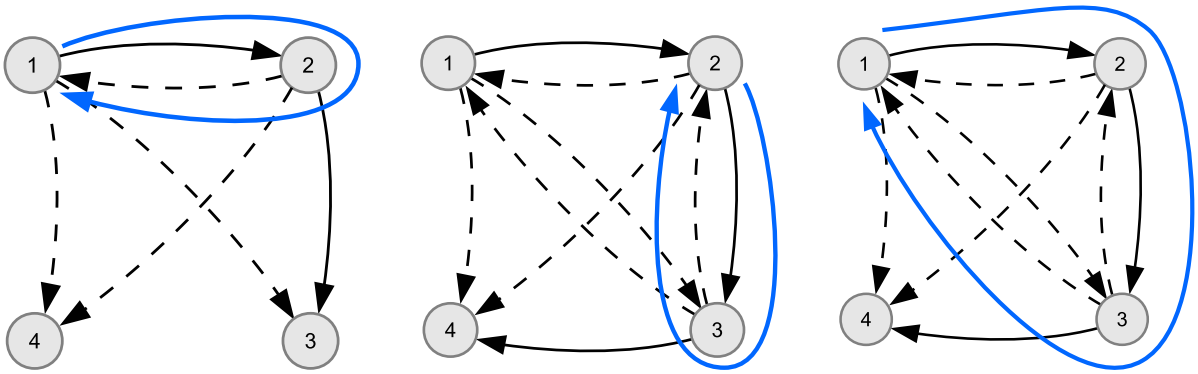
5.1 The N-Way ranging method

The N-Way ranging method is a cooperative positioning method, mainly intended to be used in wireless networks, where the localization of a large number of wireless nodes is to be performed. The N-Way ranging reduces the number of transmissions needed to obtain the positions of all nodes in a wireless network.

In the basic form of the N-Way ranging method [40], it is assumed that all nodes in the network can communicate with each other. This is usually not the case in real scenarios and not all nodes in one wireless network are visible to each other. Smaller groups of nodes, which can communicate with each other, can be selected. The N-Way ranging can be performed in these groups and the results can be further shared with the other groups. By performing additional localization between the groups, the relative locations of all nodes in the network can be estimated. If the locations of a few nodes are known, the relative locations can be easily converted to absolute. Therefore, the N-Way ranging method can be used in relative as well as in absolute positioning scenarios.

In Figure 5.1, an illustration of the N-Way ranging method is shown. In this example, only four nodes are used for the sake of simplicity. The main task is to estimate the distances between each pair of nodes. The N-Way ranging method involves N transmissions in a network where N nodes are available. The initiating node, assuming in this case to be the node 1, transmits a frame to node 2. This transmission is represented with a solid line. The other nodes can also receive this transmission and, therefore, the same frame represents a broadcast for them. Subsequently, node 2 transmits a frame to node 3 as shown in Figure 5.1a. This frame is also a broadcast to nodes 1, 3 and 4. The broadcasts are marked with dashed lines. In this case, it can be noticed that transmitting a frame from node 1 to 2 and than transmitting a broadcast from 2 to 1 is basically a TWR scenario (shown with a blue line on Figure 5.1a). Further, the node 3 transmits a frame to node 4, as shown in Figure 5.1c. This frame is also a broadcast for nodes 1, 2 and 3. At this moment, two additional distances can be estimated, the distance between the nodes 2 and 3, by using the transmission from 2 to 3 and the broadcast from 3 to 2, shown with the blue line in Figure 5.1b. This is also a typical TWR scenario. On the other hand, the distance between the nodes 1 and 3 can be estimated using the round trip time through the nodes 1, 2 and 3, as shown in Figure 5.1c. The node 1 transmits a frame to node 2, node 2 to node 3 and, finally, node 3 transmits a frame to node 4, being the same frame a broadcast to node 1. Knowing the reply times[1] as well as the time of flight (i.e distances) between nodes 1 and 2, and 2 and 3, the time of flight and, therefore, distance between the nodes 1 and 3 can be

[1] In this case, the replies are not replies in the classical sense. The **reply time** refers to the time from reception to transmission of a frame to the next node.



(a) TWR between nodes 1 and 2

(b) TWR between nodes 2 and 3

(c) Distance estimation between nodes 3 and 1

Fig. 5.1 The N-Way ranging method performed with four nodes

estimated. Finally, a frame is transmitted from node 4 to node 1. With this transmission, the N-Way ranging method is completed and the rest of the distances are estimated using the round trip times, as described previously.

As shown in Figure 5.1, for 4 nodes a total of 4 transmissions are needed. All distances between any pair of nodes can be estimated using the performed transmissions. The distance estimates are not available to all nodes after completion of NWR. Nevertheless, they can be transferred using the wireless data transmission capability of the nodes. This is also a case in the TWR approach since only the initiating node can estimate the distance.

The number of transmissions performed in the presented case is 4 for the used NWR method. If TWR is performed between each pair of nodes, the required number of transmissions would be $N(N-1) = 12$ transmissions. It is obvious from the presented case that the NWR performs significantly lower number of transmissions compared to TWR. This is especially important in cases where localization of a large number of nodes should be performed. It is also important to have a minimum number of transmission in mobile scenarios where high location update rates are needed.

The NWR method has a significantly higher implementation complexity compared to TWR. Visibility between all of the nodes performing the TWR is required. This adds additional complexity for preparation of the process. Additional data transmissions between the nodes must be performed in order to negotiate the necessary parameters needed for NWR. Anyway, some of the required required parameters are already acquired for data transmission purposes and can be reused for NWR.

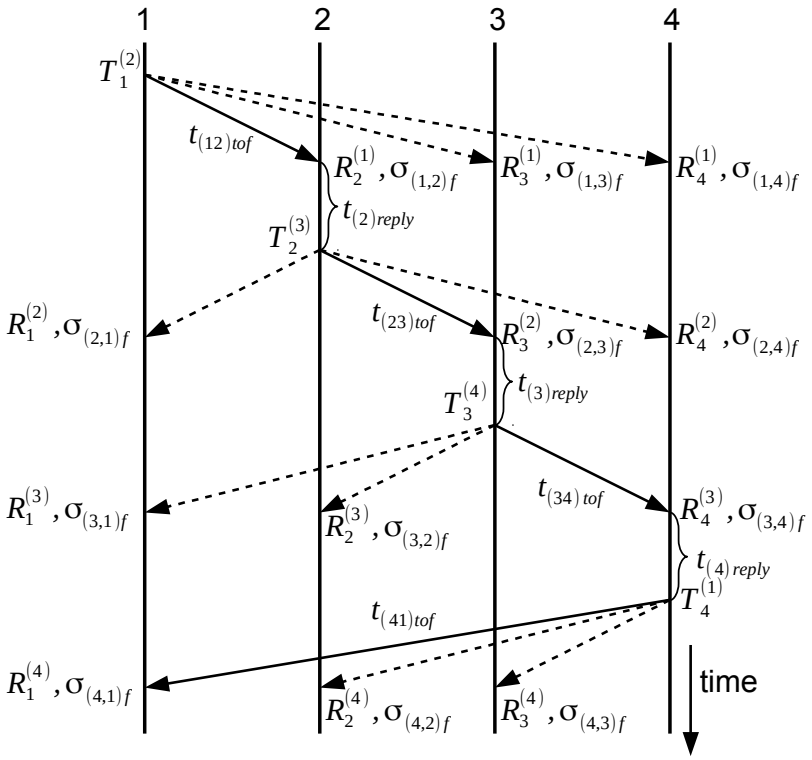


Fig. 5.2 Timing diagram of the N-Way Ranging method

5.2 N-Way ranging distance estimation

In Figure 5.2, a timing diagram of the NWR method is shown. The time instant $T_m^{(n)}$ represents the transmission time of a frame from node m to node n. The time instant $R_m^{(n)}$ represents the time of reception of a frame at node m being transmitted by node n. The time intervals $t_{(i)reply}$ represents the reply time of node i. This is the time needed for node i to transmit a frame to node $i+1$ after a frame from node $i-1$ was received. The time $t_{(mn)tof}$, is the ToF between nodes m and n. Finally, $\sigma_{(mn)f}$ represents the clock frequency offset between nodes m and n. The timing diagram of the NWR method shown in Figure 5.2 corresponds to the scenario shown in Figure 5.1. The solid arrow in the timing diagram in Figure 5.2 represent the transmission from one node to the neighboring node and the dashed arrows represent the same transmissions received as broadcasts from the remaining nodes.

All of the nodes have independent clocks used for measuring the time of reception and transmission. No synchronization between the clocks of different nodes is needed or performed. Having estimated all times of transmission, $T_m^{(n)}$, and reception, $R_m^{(n)}$, the ToF between two

neighboring nodes, i and j, where $j = i+1$, can be estimated as

$$t_{(ij)tof} = \frac{R_i^{(j)} - T_i^{(j)} - (R_j^{(i)} - T_j^{(j+1)})}{2} = \frac{t_{(i)round} - t_{(j)reply}}{2} \tag{5.2}$$

being equivalent to the TWR case. In (5.2) the round trip time at node i, $t_{(i)round}$, is estimated as $R_i^{(j)} - T_i^{(j)}$. The reply time of node j, $t_{(j)reply}$, is estimated as $(R_j^{(i)} - T_j^{(j+1)})$. Finally, the difference between the round trip time and the reply time is used to estimate the time of flight between the nodes.

When the nodes i and j are not neighboring nodes, the estimation of the distances is slightly more complex. The path for which the round trip time is estimated consists of multiple nodes. The round trip time includes multiple reply times as well as multiple times of flight. It is assumed that the times of flight between neighboring nodes are estimated in advance using (5.2). The round trip time, $R_i^{(j)} - T_i^{(j)}$, of the signal transmitted form node i, retransmitted by nodes $i+1$, $i+2\ldots$, and finally retransmitted by node j back to node i is given as

$$R_i^{(j)} - T_i^{(j)} = t_{(ij)tof} + \sum_{k=i}^{j} t_{(k,k+1)tof} + \sum_{l=i+1}^{j} t_{(l)reply} \tag{5.3}$$

where $t_{(ij)tof}$ is the ToF between nodes i and j, estimated using (5.2), $t_{(k,k+1)tof}$ is the ToF between node k and $k+1$ and $t_{(l)reply} = R_j^{(i)} - T_j^{(j+1)}$ is the reply time of the node l.

Knowing the reply times, as well as the ToFs between the neighboring nodes, the ToF between non neighboring nodes i and j $(j \neq i+1)$, $t_{(ij)tof}$, can be calculated as

$$t_{(ij)tof} = R_i^{(j)} - T_i^{(j)} - \sum_{k=i}^{j} t_{(k,k+1)tof} - \sum_{l=i+1}^{j} t_{(l)reply} \tag{5.4}$$

Using the N-Way Ranging method, the distances between each pair of nodes can be estimated using (5.2) and (5.4). Furthermore, the relative position of the nodes can be estimated using trilateration.

5.3 Ranging error due to clock frequency offset

The ranging error in the N-Way Ranging approach (and also in TWR), is a direct consequence of a long reply time in conjunction to the clock frequency offset between the nodes. In a

typical TWR ranging scenario, it is not unusual to have a reply time in the order of milliseconds. Considering two nodes, N_1 and N_2, performing TWR, the ToF estimation error can be calculated as

$$\varepsilon_{12} = \frac{1}{2} t_{(2)reply}(\sigma_{(1)f} - \sigma_{(2)f}) \tag{5.5}$$

where $t_{(2)reply}$ is the reply time of node N_2 and $\sigma_{(1)f}$ and $\sigma_{(2)f}$ are the clock frequency inaccuracies of nodes N_1 and N_2 respectively, calculated as in (3.38). Assuming a reply time of 1 ms and a clock frequency accuracy of ± 20 ppm, where one of the nodes has a clock accuracy of $+20$ ppm and the other node has -20 ppm, the ToF estimation error would be 20 ns. Taking into account that the speed of light c is 300000000 m/s, or 30 cm/ns, the ranging error $c \cdot \varepsilon_{AB}$ would be 6 meters. For precise ranging and indoor positioning applications, this magnitude of the error is unacceptable.

In the N-Way Ranging scenario this error is significantly more pronounced. Since for estimation of the distances in the N-Way Ranging method, multiple nodes have to reply sequentially to a received frame, all the reply times would be involved in the final error calculation. In cases where the clock frequency offsets are unfavorably distributed between the nodes, a huge error can be expected in the distance estimate.

The calculation of the expected error can be performed in two steps. First, the error between two neighboring nodes ($i = j+1$ or $j = i+1$)is calculated. The approach for neighboring nodes is the same as in the TWR case. In the second step, the ToF estimation error for non-neighboring nodes is calculated.

Calculation of the ranging error due to clock frequency offset for the TWR case is performed in Section 3.4.4, using (3.41), (3.42), (3.43), (3.44) and (3.45). The same approach can be used for calculation of the ranging error in the case of neighboring nodes.

Using the notation in Figure 5.2, the ToF estimation error in the case of neighboring nodes can be calculated as

$$\varepsilon_{ij} = \hat{t}_{(ij)tof} - t_{(ij)tof} = \frac{1}{2} t_{(j)reply}(\sigma_{(i)f} - \sigma_{(j)f}) \tag{5.6}$$

where $t_{(j)reply}$ is the reply time of node j, $\sigma_{(i)f}$ is the clock frequency accuracy of node i and $\sigma_{(j)f}$ is the clock frequency accuracy of node j. The difference $\sigma_{(ij)f} = \sigma_{(i)f} - \sigma_{(j)f}$ is the clock frequency offset between the nodes i and j. The hat ($\hat{}$) over the $\hat{t}_{(ij)tof}$ is used to denote that the value is an estimated value, i.e. it has the clock frequency offsets included. The $t_{(ij)tof}$ is the true value of the time of flight between the nodes.

In the case of non-neighboring nodes ($i+1 < j$), the ranging error calculation is significantly more complex. The approach is similar to the approach for neighboring nodes. The ToF between

two non-neighboring nodes i and j can be estimated as

$$\hat{t}_{(ij)tof} = (R_i^{(j)} - T_i^{(j)})(1+\sigma_{(i)f}) - \sum_{k=i+1}^{j} t_{(k)reply}(1+\sigma_{(k)f}) - \sum_{k=i}^{j-1} \hat{t}_{(k,k+1)tof} \tag{5.7}$$

where $(R_i^{(j)} - T_i^{(j)})(1+\sigma_{(i)f})$ is the round trip time measured by node i, which has clock frequency inaccuracy of $\sigma_{(i)f}$. The reply time at node k, measured with a clock having inaccuracy of $\sigma_{(k)f}$, is $t_{(k)reply}(1+\sigma_{(k)f})$ and the previously estimated ToF between two successive nodes k and $k+1$ is $\hat{t}_{(k,k+1)tof}$.

The round trip time $(R_i^{(j)} - T_i^{(j)})$ in an ideal case, measured with perfectly accurate clocks, would be

$$R_i^{(j)} - T_i^{(j)} = t_{(ij)tof} + \sum_{k=i+1}^{j} t_{(k)reply} + \sum_{k=i}^{j-1} t_{(k,k+1)tof} \tag{5.8}$$

By substituting (5.8) in (5.7), and using (3.45) for the ToF, $t_{(k,k+1)tof}$, between neighboring nodes, the final expression for $\hat{t}_{(ij)tof}$ would be

$$\hat{t}_{(ij)tof} = t_{(ij)tof} + \sum_{k=i+1}^{j} t_{(k)reply}\left(\sigma_{(i)f} - \frac{1}{2}\sigma_{(k-1)f} - \frac{1}{2}\sigma_{(k)f}\right) \tag{5.9}$$

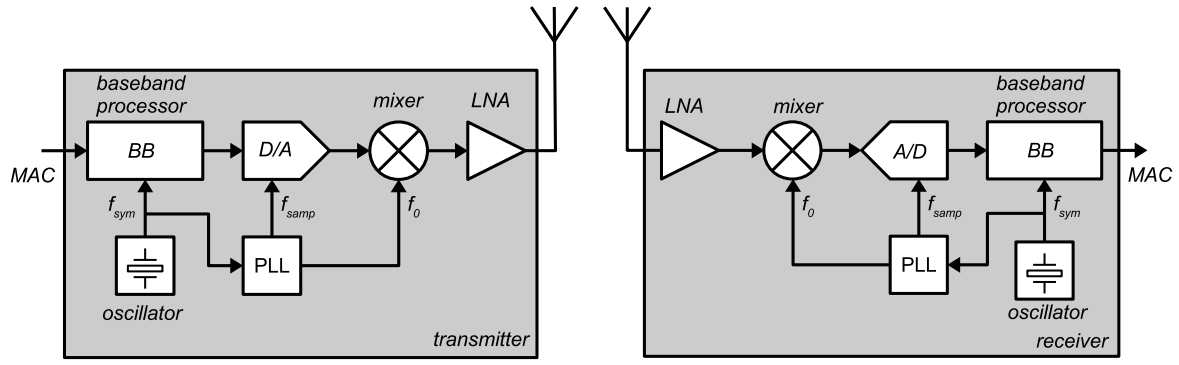
where $t_{(ij)tof}$ is the true time of flight, being $\hat{t}_{(ij)tof}$ the estimated time of flight. The remaining terms represent the error. This error can be expressed as

$$\varepsilon_{ij} = \hat{t}_{(ij)tof} - t_{(ij)tof} = \sum_{k=i+1}^{j} t_{(k)reply}\left(\sigma_{(i)f} - \frac{1}{2}\sigma_{(k-1)f} - \frac{1}{2}\sigma_{(k)f}\right), \quad for \;\; i < j-1 \tag{5.10}$$

The error calculated using (5.10) can become significant, depending on the distribution of the frequency offsets, $\sigma_{(ij)f}$, between the nodes. For example, having a network of four nodes with reply time $\Delta = 1$ *ms*, and frequency offsets of $\sigma_{(1)f} = -\sigma_{(2)f} = -\sigma_{(3)f} = -\sigma_{(4)f} = \delta = 30$ *ppm*, the produced ToF error, between nodes 1 and 4, would be

$$\varepsilon_{14} = 5\delta t_{reply} = 150 \; ns \tag{5.11}$$

In the TWR case, according to (5.5), the error would not be larger than 30 *ns*. Using the obtained ToF errors, the distance errors would be 45 meters and 9 meters in the NWR and in the TWR case respectively.



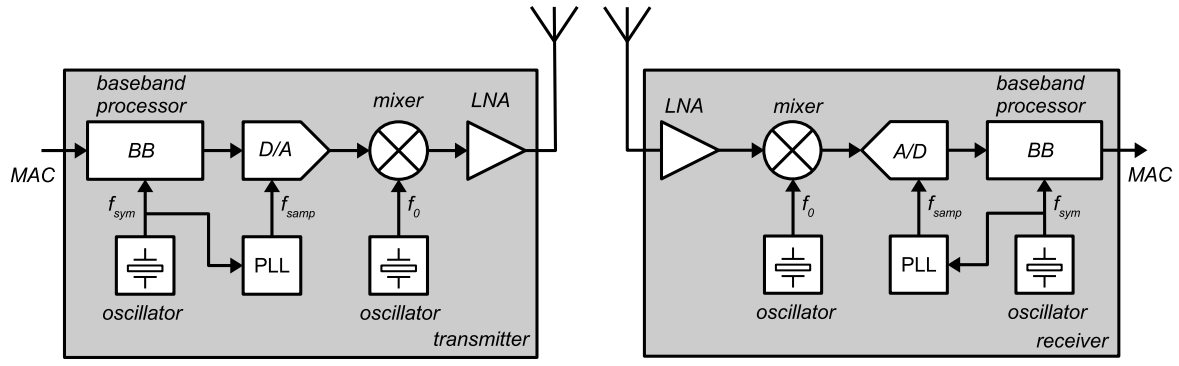
(a) Transmitter **(b)** Receiver

Fig. 5.3 Common clock used for the complete transceiver

5.4 Clock frequency offset estimation

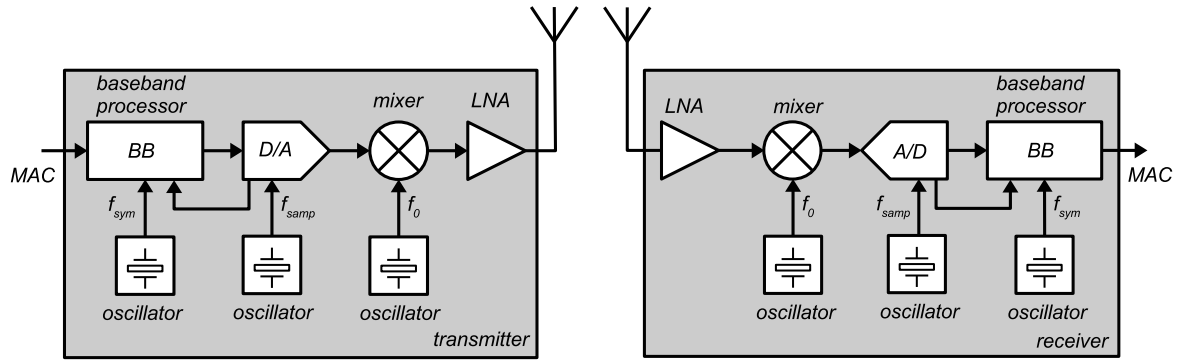
To correct the ranging error the clock frequency offset between the nodes must be estimated. This is the clock which is used for measuring the reply and round trip times when the NWR is performed. In a low end transceivers, a single clock source is used for the complete transceiver, as shown in Figure 5.3. The clocking frequencies for the baseband processor, f_{sym}, for the A/D and D/A, f_{samp} and the carrier frequency, f_0, are derived from the same clock source, using a phase locked loop (PLL). In this case, the estimated carrier frequency offset is equal to the clock frequency offset, needed to perform the ranging errors in NWR.

A more complex case is shown in Figure 5.4. In this case, the carrier frequency is derived from a separate clock source. The accuracy of this source is not the same as the accuracy of the clock used for time measurement. The carrier frequency offset estimate cannot be used for



(a) Transmitter **(b)** Receiver

Fig. 5.4 Partial separation of clock sources



(a) Transmitter **(b)** Receiver

Fig. 5.5 Complete separation of clock sources

compensation of the ranging errors. Instead, a symbol timing (i.e. symbol rate) offset should be estimated and used for compensation of the ranging error in NWR.

Finally, in the high end wireless data transceivers, the mixer, the data converters and the baseband processor usually have separate clock sources, as shown in Figure 5.5. Anyway, in order to transfer the samples between the data converters and the baseband processor, a separate clock signal, derived form f_{samp} is supplied to the baseband processor. Using this clock signal, the estimation of the clock frequency offset can be performed as in the previous case.

When quadrature receivers are used, the carrier frequency offset can be estimated by monitoring the phase of a known received signal. This known waveform can be a preamble of the incoming frame as in [109]. Another option is to use a known waveform (frame or sequence) only for the purpose of ranging and frequency offset estimation, as in [96]. Many other carrier frequency offset estimation techniques are available [16, 82, 100, 114] and can be used for the purpose of correction of the distance estimation errors due to clock frequency offset in NWR. If separate clock sources are used for the carrier frequency and baseband processor, the symbol timing should be estimated. Many receivers perform symbol timing estimation and a large amount of algorithms are available [37, 80].

In the proposed approach it is assumed that the clock frequency offset estimation is performed each time a frame is received. In the scenario presented in Figure 5.2, when node i transmits a frame, the remaining nodes estimate the offset $\sigma_{(ij)f} = \sigma_{(i)f} - \sigma_{(j)f}$ between themselves and node i. As stated in [13], the carrier signal can be typically tracked with a precision of 30° ($\frac{1}{12}$ of a cycle). It can be assumed that once the crystal clock frequency offset is estimated it will remain stable during the ranging process. The main reason for this claim is that the temperature and the age of the crystal[2] do not change rapidly (in the order of *ms*).

[2]Temperature change as well as aging can introduce significant ToF errors.

5.5 Clock frequency offset compensation

In the proposed approach the estimated clock frequency offset is used to correct the estimated ToF and, therefore, the estimated distance error. Since the clock frequency offset can be precisely estimated, the correction can be applied and the distance estimation error would be significantly reduced.

The starting point for estimation of the distances are the measured round trip times, $\widetilde{t}_{(ij)round}$, as well as the measured reply times, $\widetilde{t}_{(j)reply}$. The tildes ($\widetilde{\ }$) represent time intervals measured using the inaccurate clock sources present at the nodes. Node i measures the round trip time, while nodes j measure the reply times. The clock frequency of node i has an frequency inaccuracy of $\sigma_{(i)f}$ and the clock frequencies of nodes j have a frequency inaccuracies of $\sigma_{(j)f}$. If the measured round trip times and reply times are used for ToF estimation, the obtained range estimates would include the errors due to clock frequency offset.

In order to correct the ToF measurements, correction factors should be calculated and applied using the measured reply times as well as the estimated clock frequency offsets between the nodes. Again, the problem is addressed in two steps. In the first step, the correction factor for the ToF between two neighboring nodes ($i+1=j$) and, in the second step, the factors for correction of the ToF between non-neighboring nodes ($i+1<j$) is calculated. There are several reasons for solving the problem in two steps. The case with two neighboring nodes is equivalent to the TWR approach. The obtained correction factors can be used in NWR as well as in TWR. On the other hand, the case with non neighboring nodes is more general and, at the same time, more complicated. It can be probably adapted for use in both cases (neighboring and non neighboring nodes), but that would add to its complexity, making it inefficient for practical use.

In order to estimate the ToF in the case of two neighboring nodes, the round trip time should be measured first. Knowing the measured time of transmission of the frame, $\widetilde{T}_i^{(j)}$, as well as the measured time of reception of the reply, $\widetilde{R}_i^{(j)}$, as shown in Figure 5.2, the round trip time can be estimated as

$$\widetilde{t}_{(ij)round} = \widetilde{R}_i^{(j)} - \widetilde{T}_i^{(j)}, \tag{5.12}$$

The time interval $\widetilde{t}_{(ij)round}$ is the measured round trip time. A relation between the measured and the true, $t_{(ij)round}$, would be

$$\widetilde{t}_{(ij)round} = \widetilde{R}_i^{(j)} - \widetilde{T}_i^{(j)} = t_{(ij)round}(1+\sigma_{(i)f}). \tag{5.13}$$

The true round trip time can be also written as

$$t_{(ij)round} = 2t_{(ij)tof} + t_{(j)reply} = 2t_{(ij)tof} + \widetilde{t}_{(j)reply}(1 - \sigma_{(j)f}), \tag{5.14}$$

The round trip time, in (5.14) is consisted of two ToFs, i to j and back j to i, plus the reply time $t_{(j)reply}$.

In (5.14) the $t_{(j)reply}$ value is approximated as $\widetilde{t}_{(j)reply}(1 - \sigma_{(j)f})$. The approximation is performed starting from the equation

$$\widetilde{t}_{(j)reply} = t_{(j)reply}(1 + \sigma_{(j)f}) \tag{5.15}$$

where $\widetilde{t}_{(j)reply}$ is the measured reply time, $t_{(j)reply}$ is the true reply time and $\sigma_{(j)f}$ is the frequency inaccuracy of the clock source used for measuring the reply time.

The true reply time would therefore be

$$t_{(j)reply} = \frac{\widetilde{t}_{(j)reply}}{(1 + \sigma_{(j)f})} \tag{5.16}$$

Using the Taylor series, and the first two elements form the series, the reply time would finally be

$$t_{(j)reply} \approx \widetilde{t}_{(j)reply}(1 - \sigma_{(j)f}) \tag{5.17}$$

Further, substituting (5.14) in (5.13), the measured round trip time is expressed as

$$\widetilde{t}_{(ij)round} = (2t_{(ij)tof} + \widetilde{t}_{(j)reply}(1 - \sigma_{(j)f}))(1 + \sigma_{(i)f}). \tag{5.18}$$

The node i, in order to estimate the ToF, measures the round trip time $\widetilde{t}_{(ij)round}$. The reply time of the node j is either negotiated in advance or is simply measured by node j and sent back to node i. Node i further estimates the ToF as

$$\hat{t}_{(ij)tof} = \frac{\widetilde{t}_{(ij)round} - \widetilde{t}_{(j)reply}}{2}. \tag{5.19}$$

By substituting (5.18) in (5.19) we get

$$\hat{t}_{(ij)tof} = \frac{(2t_{(ij)tof} + \widetilde{t}_{(j)reply}(1 - \sigma_{(2)f}))(1 + \sigma_{(1)f}) - \widetilde{t}_{(j)reply}}{2}, \tag{5.20}$$

By rearranging and simplifying (5.20) the ToF can be estimated as

$$\hat{t}_{(ij)tof} = t_{(ij)tof}(1+\sigma_{(1)f}) + \frac{1}{2}\widetilde{t}_{(j)reply}(\sigma_{(1)f} - \sigma_{(2)f} - \sigma_{(1)f}\sigma_{(2)f}) \tag{5.21}$$

Since $\sigma_{(i)f} \ll 1$ and $\sigma_{(j)f} \ll 1$, the factor $\sigma_{(i)f}\sigma_{(j)f}$ can be assumed to be negligible compared to the sum $\sigma_{(i)f} - \sigma_{(j)f}$. An approximation $1+\sigma_{(i)f} \approx 1$ can also be performed. Finally, the estimated time of flight is represented by

$$\hat{t}_{(ij)tof} = t_{(ij)tof} + \frac{1}{2}\widetilde{t}_{(j)reply}(\sigma_{(i)f} - \sigma_{(j)f}), \tag{5.22}$$

where the factor $\frac{1}{2}\widetilde{t}_{(j)reply}(\sigma_{(i)f} - \sigma_{(j)f})$ is the error due to the clock frequency offset between nodes i and j. Subtracting the error from (5.20) would compensate the error due to clock frequency offset. The final equation for compensation of the estimated ToF would therefore be

$$\bar{t}_{(ij)tof} = \frac{1}{2}(\widetilde{t}_{(ij)round} - \widetilde{t}_{(j)reply}(1+\sigma_{(i)f} - \sigma_{(j)f})). \tag{5.23}$$

The value $\sigma_{(i)f} - \sigma_{(j)f}$ in (5.23) is the estimated clock frequency offset between nodes i and j. The bar $(\bar{\cdot})$ over the estimated ToF means that the error due to clock frequency offset is compensated.

For non neighboring nodes, $t_{(ij)round}$, (5.24) represents the round trip time for the frame transmitted from node i received and retransmitted from nodes $i+1, i+2, ..., j$ sequentially, and finally received back by node i. An example of this is shown in Figure 5.1c.

The true round trip time (without offset errors included), in an ideal case, is

$$t_{(ij)round} = t_{(ij)tof} + \sum_{k=i}^{j-1} t_{(k,k+1)tof} + \sum_{k=i+1}^{j} t_{(k)reply}. \tag{5.24}$$

On the other hand, $\hat{t}_{(ij)tof}$ can be estimated as

$$\hat{t}_{(ij)tof} = \widetilde{t}_{(ij)round} - \sum_{k=i}^{k=j-1} \bar{t}_{(k,k+1)tof} - \sum_{l=i+1}^{j} \widetilde{t}_{(l)reply}, \tag{5.25}$$

using the measured values for the round trip time, $\widetilde{t}_{(ij)round}$, and the reply time, $\widetilde{t}_{(l)reply}$. The symbol $\bar{t}_{(k,k+1j)tof}$ in (5.25) represents the clock frequency offset error compensated version of the $t_{(k,k+1)tof}$, from (5.23), and can be assumed equal to it. Further, (5.25) is derived from (5.24), using the measured values instead of the true values. The measured values do not

include clock frequency offset compensation and, therefore, they would introduce additional large errors in the ToF estimation.

The measured round trip time, $\widetilde{t}_{(ij)round}$, with the error due to the clock frequency offset included, can be expressed as

$$\widetilde{t}_{(ij)round} = R_i^{(j)} - T_i^{(j)} = t_{(ij)round}(1+\sigma_{(i)f}). \tag{5.26}$$

Substituting $t_{(j)reply}$ from (5.17) in (5.24), would express the $t_{(ij)round}$ as a function of the measured reply time $\widetilde{t}_{(k)reply}$, as well as the clock frequency offset $\sigma_{(k)f}$. The obtained equation would be

$$t_{(ij)round} = t_{(ij)tof} + \sum_{k=i}^{j-1} t_{(k,k+1)tof} + \sum_{k=i+1}^{j} \widetilde{t}_{(k)reply}(1-\sigma_{(k)f}). \tag{5.27}$$

Further substituting (5.27) in (5.26) gives

$$\widetilde{t}_{(ij)round} = t_{(ij)tof}(1+\sigma_{(i)f}) + \sum_{k=i}^{j-1} t_{(k,k+1)tof}(1+\sigma_{(i)f}) + \sum_{k=i+1}^{j} \widetilde{t}_{(k)reply}(1-\sigma_{(k)f}+\sigma_{(i)f}). \tag{5.28}$$

Finally, substituting $\widetilde{t}_{(ij)round}$ from (5.28) in (5.25), leads to

$$\hat{t}_{(ij)tof} = t_{(ij)tof} + \sum_{k=i+1}^{j} \widetilde{t}_{(k)reply}(\sigma_{(i)f}-\sigma_{(k)f}). \tag{5.29}$$

The sum in (5.29) is the ranging error due to the clock frequency offset. To compensate the ToF, this error should be subtracted from (5.25). The final equation for the corrected ToF, based on the measured values would therefore be

$$\bar{t}_{(ij)tof} = \widetilde{t}_{(ij)round} - \sum_{k=i}^{k=j-1} \bar{t}_{(k,k+1)tof} - \sum_{l=i}^{j-1} \widetilde{t}_{(l)reply} - \sum_{k=i+1}^{j} \widetilde{t}_{(k)reply}(\sigma_{(i)f}-\sigma_{(k)f}). \tag{5.30}$$

In (5.30) the error due to clock frequency offset is compensated.

5.6 Numerical simulation and results

The performance of the proposed method is assessed in simulation. The NWR method simulation results, without compensation of the ranging error, are compared with the results obtained

using the proposed approach. Multiple simulations are performed for the given scenarios in order to obtain statistics.

To perform the simulation, a model of the system has been developed. Each node is modeled with its reply time and clock frequency inaccuracy. The simulation is performed for a square two dimensional area with a length and width of d. The nodes are positioned randomly in this area. The number of nodes, N, is varied in order to evaluate the ranging error as a function of number of nodes. Also, simulations for different values of the clock frequency inaccuracy are performed.

The N-Way ranging is performed and the corresponding frequency offset is included in each measured reply and round trip time. The ToF between each pair of nodes is estimated using the measured values. To compare the proposed approach, a simulation with compensation of the ToF estimates, is performed. The same position of the nodes, as in the uncompensated case are used.

The simulations are performed using 3 to 10 nodes, located randomly within an area of 10 x 10 meters. The crystal clock frequency inaccuracy is selected randomly for each node within the specified limits (in this case 10, 20, 30, 40 and 50 ppm). The clock frequency offset estimation error is randomly selected within 10 percent of the true value. In order to obtain good statistical results, for each number of nodes and for each clock frequency inaccuracy, a total of 100 simulations are performed and the obtained results averaged.

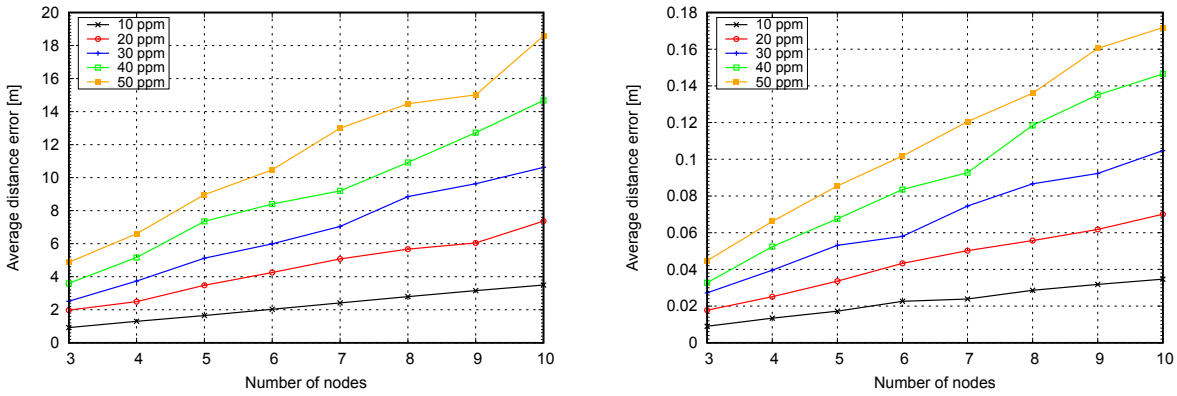
It should be noted that in these simulations only the errors due to clock frequency offset are of interest. Therefore, only these errors are modeled and obtained in simulation. Other sources of errors are not included in the models used, since they would introduce additional errors in the distance measurements.

The root mean square error (RMSE) is calculated as

$$\varepsilon_{RMSE} = \sqrt{\frac{1}{N(N-1)/2} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \varepsilon_{ij}} \tag{5.31}$$

where ε_{RMSE} is the RMSE of the estimated distances, in a given simulation scenario, and ε_{ij} is the estimation error of the distance between the nodes i and j.

In Fig. 5.6a, measurement errors for the uncompensated Multi-Way ranging are presented. On the x-axis the number of nodes for each simulation is shown. The y-axis represents the average RMSE after performing 100 simulations. The 5 different curves in the plot are for different clock frequency inaccuracies (10 - 50 ppm). The error increases with the number of nodes, since it accumulates from larger number of nodes. As can be seen from Fig. 5.6a,



(a) Distance RMSE without compensation

(b) Distance RMSE with compensation

Fig. 5.6 Distance RMSE obtained by simulation of the NWR method without and with ranging error compensation

for clock frequency inaccuracy larger than 20 *ppm* and more than 5 nodes, the distance error becomes significant (± 5 m), compared to the distances between the nodes. Thus, without any clock frequency offset compensation, it is not possible to perform Multi-Way ranging with a reasonably small distance error.

The second set of simulations include the proposed clock frequency offset compensation method, using the same simulation parameters. The results in Fig. 5.6b show the same trend, except that the error in the worst case is less than 0.18 *m*. In contrast to this, the error in the first simulation, Fig. 5.6a, was larger than 18 *m*.

The results show that the proposed method significantly reduces the ranging error due to clock frequency offset for the NWR method.

5.7 Comparison with other approaches for compensation of the ranging error due to clock frequency offset

In this section, the proposed approach is compared to the TWR and NWR methods used in cooperative localization scenarios. In the TWR case, two different compensation approaches can be used: compensation of ranging error using the estimated clock frequency offset between the nodes and the Symmetrical Double Sided TWR (SDS-TWR) [43]. The performances of both approaches, when used for ranging in cooperative localization scenarios, are presented. The Symmetric N-Way Ranging is also compared with the proposed approach. It compensates the ranging error due to clock frequency offset between the neighboring nodes, but doubles the number of transmissions.

Table 5.1 Number of transmissions needed for localization of n nodes and average ranging error for different approaches

Localization method	No. of transmissions	ranging error
Two Way Ranging	$\approx n^2$	$\frac{1}{2}t_{reply}\sigma_{(NN)f}$
N-Way Ranging	n	$\frac{1}{2}Nt_{reply}\sigma_{(NN)f}$
Compensated Two Way Ranging	$\approx n^2$	$\frac{1}{2}t_{reply}\sigma_{(NN)f}\varepsilon_{(NN)f}$
SDS-TWR between each pair of nodes	$\approx \frac{3}{2}n^2$	$t_{tof}\sigma_f$
Symmetrical N-Way Ranging	$2n$	$\frac{1}{2}Nt_{reply}\sigma_{(NN)f}$
Proposed approach in this work	n	$\frac{1}{2}Nt_{reply}\sigma_{(NN)f}\varepsilon_f$

The average error is calculated and presented. The results are summarized in Table 5.1. The details for derivation of these errors are given in Appendix A.

The main advantage of the proposed method, over the others, is the minimal number of transmissions and excellent ranging error compensation. Moreover, it halves the number of transmissions required compared to the Symmetric N-Way method. Therefore, it saves energy and reduces the medium usage. If the proposed method is implemented on a typical wireless data transceiver, it would not considerably increase its complexity. The required clock frequency offset estimation is usually performed in a data transceiver and, therefore, it can be reused for localization.

Chapter 6

Implementation of ranging and positioning methods

The proposed METS approach in Chapter 4 and the proposed approach for compensation of the ranging error due to clock frequency offset in Chapter 5, require a data transmission capability of nodes for exchanging the parameters as well as estimated distances and estimated times. Therefore, in this chapter an approach for implementing ranging methods on a wireless node and their integration with data transmission functionality is presented. The proposed approach is used for implementation of TWR on a software defined radio (SDR) platform. It was tested in the 2.4 and 5 GHz ISM bands. Additionally, an integration of ranging and data communication was performed on a 60 GHz system. The 60 GHz system is preferred for integration of both functionalities, due to the availability of hardware resources as well as intellectual property (IP) cores for data transmission and ranging. Nevertheless, the hardware capabilities of the used SDR are quite limited, making the integration of both functionalities almost impossible.

6.1 Ranging and positioning in the 2.4/5 GHz band using software defined radio system

The ranging and positioning system was implemented on an Ettus Research SDR system[1]. The used USRP N210 model SDR consists of a few main parts as shown in Figure 6.1. The initial signal processing is performed in the Field Programmable Gate Array (FPGA). The SDR is connected to a host computer using Gigabit Ethernet connection. The signal processing

[1] http://www.ettus.com

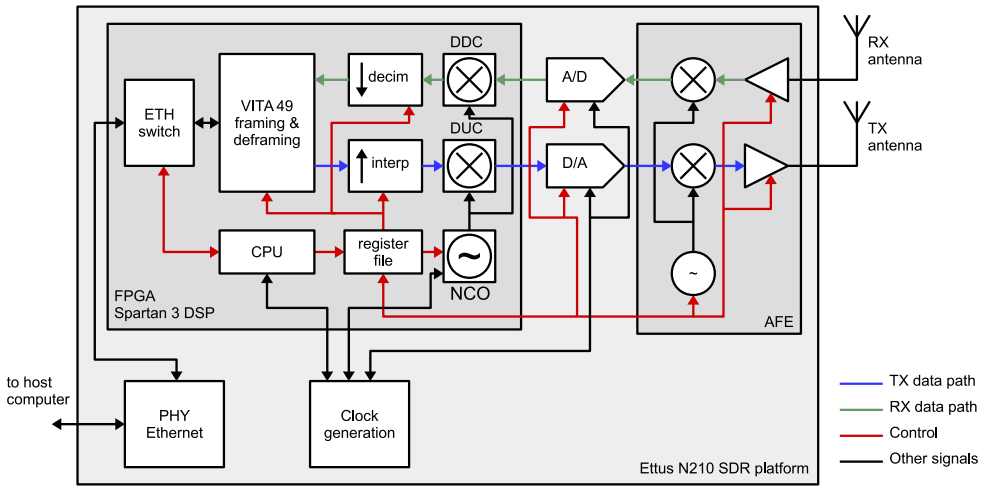


Fig. 6.1 Simplified block diagram of a USRP N210 software defined radio platform

Table 6.1 The main specifications of the USRP N210 SDR

Characteristic	Value	Unit
ADC resolution	14	bits
ADC sample rate	100	MSps
ADC wideband SFDR	88	dBc
DAC resolution	16	bits
DAC sample rate	100	Msps
DAC wideband SFDR	80	Msps
Ethernet streaming sample rate	50 (@ 8 bits/sample)	Msps
	25 (@ 16 bits/sample)	Msps
TCXO frequency reference accuracy	2.5	ppm
with GPSDO	0.01	ppm
FPGA	Spartan 3A-DSP 3400 FPGA (N210)	
High speed RAM	1	MB

required by the user is performed on the host computer. The baseband complex samples are transferred to/from the host system using the Gigabit Ethernet connection.

The main characteristics of the USRP N210 are summarized in Table 6.1

An Analog Front-End (AFE) is used for up/down conversion of the baseband signals. The analog baseband signals needed for the AFE are converted using A/D and D/A converters. The analog front-end is a separate daughter card inserted into the SDR. There are a few different cards that can be used, and they differ mainly in the frequency band they cover. For this implementation the daughter card XCVR2450, also from Ettus Research, was used. The basic specs are given in Table 6.2

Table 6.2 The main specifications of the XCVR2450 daughter card

Characteristic	Value	Unit
Frequency band		
2.4 GHz	2.4 - 2.5	GHz
5 GHz	4.9 - 5.9	GHz
Output power	100	mW
Max RX bandwidth	36	MHz
Max TX bandwidth	36	MHz

This daughter card is selected since it covers the ISM bands at 2.4 and 5 GHz[2], for which no licensing is needed. For the performed tests, the 5 GHz band was preferred, due to the lower interference from other devices utilizing ISM bands.

The FPGA, as shown in Figure 6.1, performs minimal signal processing on the transmit and receive signals. There are two main signal processing units shown in Figure 6.1. The first block contains the interpolator for the transmit signal and the decimator for the received signal. They actually adapt the data rate of the data converters, with the data rate of the samples from the host system connected to the Gigabit Ethernet port. Both, A/D and D/A, converters acquire samples at a fixed sample rate of 100 MSps, which is decimated down to the rate required by the host. Anyway, the maximum sample rate is limited by the Gigabit Ethernet connection to 25 MSps for 16 bit wide complex samples, or 50 MSps for 8 bit wide complex samples. The anti-aliasing and interpolation filters are included in the decimators and interpolators.

The Digital Up-Converter (DUC) and the Digital Down-Converter (DDC), shown in Figure 6.1 are used for fine tuning of the carrier frequency. The AFE is only capable of setting a finite number of discrete carrier frequencies. The tuning resolution of the AFE is usually not fine enough for many applications and, therefore, the final carrier frequency is tuned in two steps. The coarse tuning of the carrier frequency is performed by the AFE and the additional fine tuning is performed using the DUC and DDC. A resolution of under 1 Hz can be easily achieved with this approach.

[2]The frequency bands are 2.400-2.500 GHz and 5.725-5.875 GHz.

6.1.1 Two way ranging implementation on a software defined radio platform

Among the available localization methods we have selected the TWR for implementation in the SDR platform due to its relaxed implementation requirements. The TWR approach was previously described and is shown in Figure 3.2b.

The TWR can be implemented in software as well as in the hardware of the available SDR platform. The software implementation is possible, due to the time stamps associated with the complex samples arriving from and going to the SDR platform. The main issue in this approach is the long reply time. The complex samples are transferred using Gigabit Ethernet connection to the host computer, in the SDR platform described previously. Since the transmit and receive samples should be packed in UDP packets and sent using Ethernet frames, additional delay would be introduced in the reply time.

An additional problem is the available sample rate. The data converters in the used SDR platform have a sampling rate of 100 MSps, which cannot be transfered to the host computer, due to limited bandwidth of the Gigabit Ethernet. The maximum achievable sample rate at the host computer is 50 MSps. For ranging purposes it is important to have the highest possible sample rate, available in the SDR, in order to perform a precise distance estimation.

Due to the above stated issues, the implementation is partially performed in hardware and partially in software. The time critical tasks are performed in hardware, while the non time critical tasks are performed in software. The TWR process itself is implemented in hardware. The complex samples acquired during TWR are stored in the RAM of the SDR platform. Later, these samples are transferred to the host platform, where the distance estimation is performed in software.

Using this approach different distance estimation algorithms can be tested. Implementation of distance estimation algorithms in hardware would be possible but, for testing purposes, this is not justified, given the implementation effort. In this implementation, data transmission is not of interest and, therefore, it is not implemented.

A slight difference in the TWR is introduced for this implementation. Usually, when two nodes perform TWR, they transmit frames locally generated in each node. In the implemented approach the node N_1, shown in Figure 3.2b, generates and transmits a frame to node N_2. The node N_2 receives this frame and saves the acquired samples in memory. Further, these samples are retransmitted back to node N_1 as a reply. This approach is also used in [83]. The advantage of using this approach is that the effects in the received signal at node N_1 due to carrier frequency offset would be nullified.

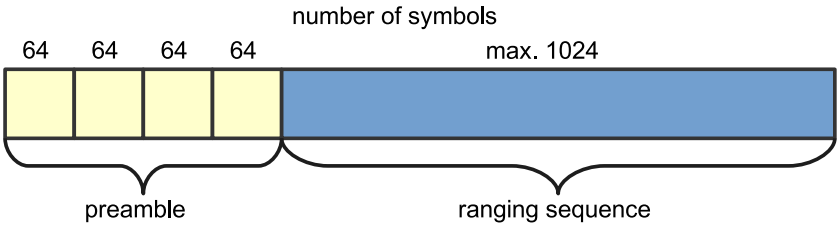


Fig. 6.2 Frame format used for performing TWR implemented on a SDR platform

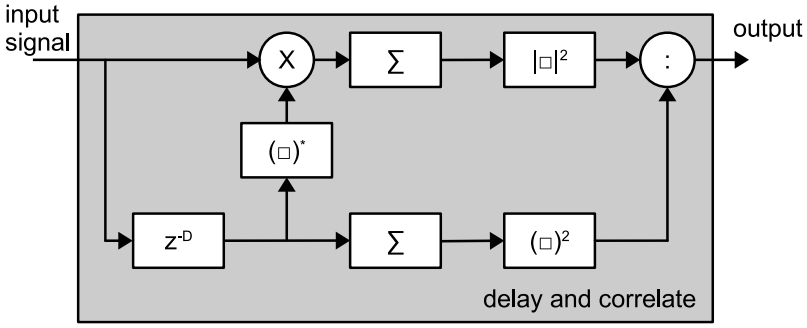


Fig. 6.3 Block diagram of a ***delay and correlate*** preamble detector

Anyway, reception and retransmission of the same signal, without reconstruction of the received symbols, leads to lower SNR at node N_1. The signal received at node N_2 contains noise. Node N_2 retransmits the received signal as well as the contained noise. Additional noise would be added when this signal is received back in node N_1. Since the system is used for testing purposes, a simple solution to this problem is to increase the transmit powers at nodes N_1 and N_2, in order to increase the SNR. This is not the optimal solution, but it is less complex compared to the solution involving carrier frequency offset estimation and compensation at node N_1. This approach would reduce the usable distance, but this is not a significant disadvantage, since the developed system is used only for testing purposes.

The frame structure and preamble detection

The structure of the frame used for ranging is shown in Figure 6.2. This frame is created at the host computer and transferred to the RAM of the SDR using the Gigabit Ethernet connection. A preamble is placed at the beginning of the frame. It consists of four copies of a short PN sequence with a length of 64, modulated using BPSK modulation. This preamble is used by the receiver to detect the frame arrival and it has a structure that can be easily detected using an autocorrelator or a so called "delay and correlate" preamble detector [46]. Its block diagram is shown in Figure 6.3. The input waveform $x[n]$ is multiplied by its delayed (z^{-D}) and conjugated

$((\cdot)^*)$ version. The product is then fed into a moving average filter, S_1. The delay factor, D, is equal to the length of the PN sequence used in the preamble, in this case 64. The second moving average filter computes the average power of the incoming signal. When a preamble is received, the signal $a[n]$ would present a high value due to the high correlation between the received signal and the delayed version of it. The average power is used as a normalization factor allowing reliable operation of the preamble detector in both low and high SNR scenarios.

The detection of the preamble time of arrival, which triggers the sample acquisition at node N_2, is not precise. A part at the beginning of the preamble can be lost, due to late detection of the incoming frame. If the preamble consists of only two modulated PN sequence copies instead of four, one copy would not be fully acquired at node N_2. In this case, there would be no two identical successive copies in the preamble of the frame transmitted by node N_2 and frame detection at node N_1 would be complicated. Therefore, we include four copies of the modulated PN sequence in the preamble to ensure reliable preamble and frame detection at node N_1.

The second part of the frame is the ranging sequence, as shown in Figure 6.2. Different PN sequence types can be used for ranging, including m-sequences, Gold sequences, Kasami etc. In this implementation the ranging sequence is also BPSK modulated.

6.1.2 The hardware TWR core

As shown in Figure 6.1, the SDR consists of one FPGA to which the data converters and the Ethernet physical layer chip are connected. The hardware description language (HDL) code for configuration of the FPGA is open source and is available to the user. The host computer driver is also open source and free for modifications.

The real-time implementation of the TWR method is performed on the FPGA. The range estimation as well as the positioning is implemented in software on the host computer. Trilateration is used for position estimation.

A modification of the control logic, as well as, changes in the transmit and the receive path of the SDR FPGA configuration are introduced in order to implement the TWR method. In the control logic, functionalities for controlling of the ranging process must be implemented. It should control the transmission and reception of the ranging frame, detection of the frame arrival, as well as the communication with the host computer. The coarse measurement of the round trip time must be implemented in hardware, since the reply time is fixed and its precise measurements in software is challenging. Memory buffers for transmission and reception of the

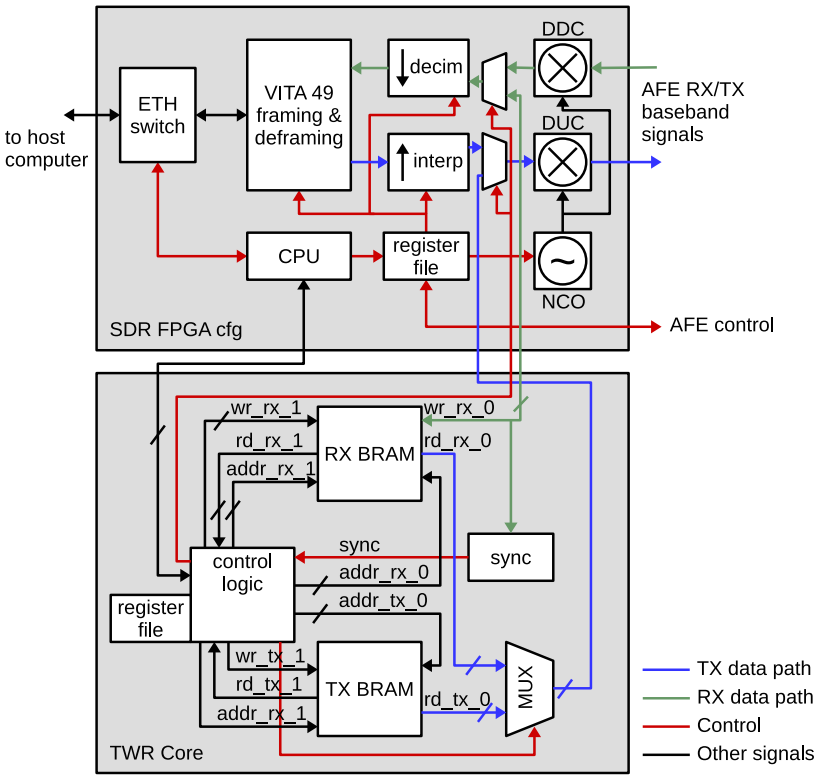


Fig. 6.4 Block diagram of the SDR FGPA configuration with the TWR core included

ranging waveforms by the ranging system must be implemented. A synchronizer, for detection of frame arrival is also implemented in the receive path.

A block diagram of the SDR FPGA configuration with the modifications included is shown in Figure 6.4. The "SDR FPGA cfg" block is the original SDR FPGA configuration of the SDR. It is similar to the block diagram of the FPGA configuration shown in Figure 6.1. The main difference in the "SDR FPGA cfg" block is the presence of an additional multiplexer and the demultiplexer. Their purpose is to route the signal either from/to the host computer or from/to the ranging core i.e. "TWR Core". Also, a connection to the soft core CPU[3] is added. This connection is used to write and read parameters in the register file of the **TWR core** as well as to access the signal samples stored in the RAMs of the core.

The ranging core i.e. ***TWR Core*** has a few basic blocks. The process of ranging is controlled by the ***control logic*** block. The parameters needed for the process, as well as the results and status flags for the TWR process are stored in the ***register file***. The block ***sync*** is a ***delay and***

[3] A soft core CPU is a CPU which is implemented on a programmable hardware, like FPGA. The performances of such CPU are expected to be lower compared to a hardware CPU. Anyway, for many applications the performances of the soft core CPU are more than satisfying. The flexibility of changing the configuration of the CPU is additional advantage.

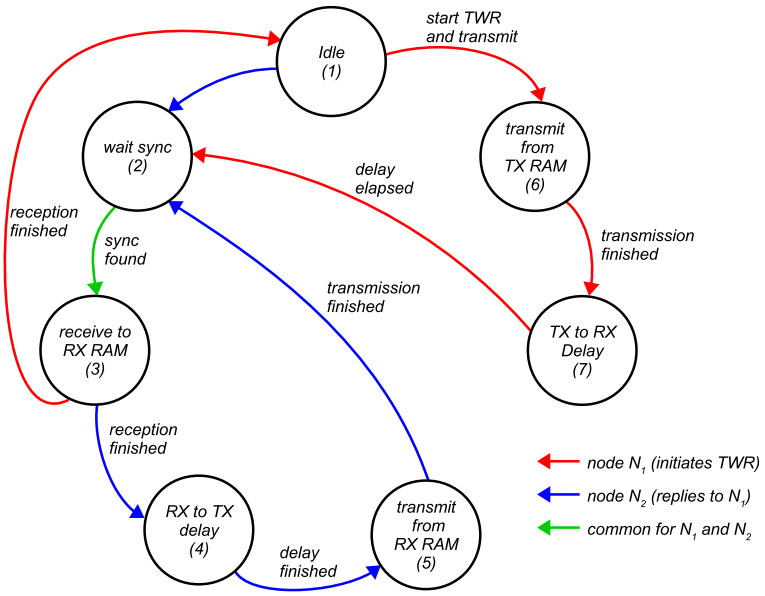


Fig. 6.5 Finite state machine diagram of the control block

correlate preamble detector. There are two memories in the ranging core referred to as ***RX BRAM*** and ***TX BRAM***. These memories are so called Block RAM, a type of RAM memory integrated in the FPGA [116]. These memories store the samples that should be transmitted and the received. The multiplexer ***MUX*** is used to route the samples from the BRAMs to the D/A converter, i.e. the transmitter, depending on the role of the node.

Before the ranging is started, the host computer configures the nodes. At first, the role of the nodes is configured. One of the nodes is configured as an initiating node, N_1 in Figure 3.2b, and the second node is configured as a reflecting node, N_2 in Figure 3.2b. Further, the frame, shown in Figure 6.2 is uploaded, from the host computer, in the ***TX BRAM*** of the initiating node N_1. The reflecting node N_2, on the other hand, is set in a listening i.e. reflecting mode, by the host computer.

In order to perform the TWR, the host computer sends a signal to the initiating node , to start the TWR process. The control of the TWR process is performed by the ***control logic*** block, shown in Figure 6.4. This block is implemented as an finite state machine (FSM) and its state diagram is given in Figure 6.5. The FSM is in state *1 (idle)* when it receives the command to start TWR. It configures the multiplexers accordingly, in order to transmit the samples from the ***TX BRAM***, activates the transmitter, and goes to state *6 (transmit from TX RAM)*. The samples from the ***TX BRAM***, containing the ranging frame, are transmitted. When the transmission is finished, the transmitter is turned off and the receiver is turned on and the FMS goes into state *7 (TX to RX delay)*. When this short delay elapses, the FSM goes into state

2 (wait sync) where it waits for the reply from the reflecting node N_2.

The ***control logic*** block in the reflecting node N_2, controls the node functionality after being initialized by the host computer and its FSM goes into state *2 (wait sync)*. When the preamble is detected by the preamble detector, ***sync***, the FSM goes into state *3 (receive to RX RAM)* and stores the received signal samples in ***RX BRAM***. The number of samples that are stored in ***RX BRAM*** is predetermined by the host computer. When the required number of samples is acquired in the ***RX BRAM*** a short delay is performed in state *4 (RX to TX delay)* and the FSM goes into state *5 (transmit from RX RAM)*. It starts retransmitting the samples from the ***RX BRAM***, back to the initiating node N_1. These samples include at least half of the preamble and the whole ranging PN sequence, BPSK modulated. When the transmission is finished, the FSM goes into state *2 (wait sync)*, where it waits for the next incoming frame.

In the meantime, the originating node is switched to receiving mode, and waits for the reception of the reflected frame in state *2 (wait sync)*. When the preamble is detected, the node acquires a predetermined number of samples in the ***RX BRAM***, in the state *3 (receive to RX RAM)*. The length of the time interval from the start of transmission to the start of reception is measured and stored in the register file. This is the value of the coarse round trip time. When the required number of samples are acquired, the FSM goes into state *1 (idle)*. The samples from the ***RX BRAM*** are later transferred to the host computer in order to perform fine ToF estimation. Also the coarse round trip time value is transferred to the host computer. Without it the ToF estimation cannot be performed.

6.1.3 Distance estimation

Distance estimation is implemented in software on the host computer. This allows the user to easily change different parameters of the estimator, which is not the case if they were implemented in hardware. The software implementation is slower compared to hardware implementation, but for evaluation purposes the speed is not a limiting factor. The qualitative parameters, like the ranging precision of different estimators, are more important.

An estimator, like the one shown in Figure 3.22 is used for ToA estimation of the frame arriving at node N_1 and being transmitted from node N_2. The round trip time is estimated using the ToA and the time of transmission. The reply time at node N_2 is fixed and must not be estimated. Knowing both reply and round trip time, the ToF is finally estimated. Spline interpolation between the samples is used to improve the ranging precision.

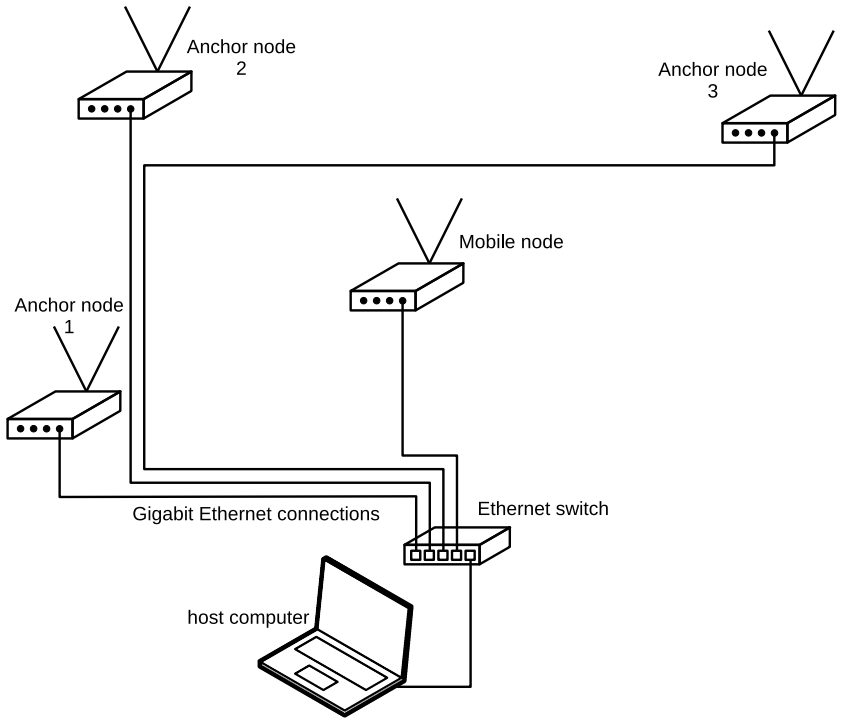


Fig. 6.6 Physical architecture of the developed positioning system

6.2 Architecture of the localization system implemented on a SDR platform

Localization of a node can be performed using at least three distance estimates to static anchor nodes with known position. The physical architecture of the localization system which is developed in this work is shown in Figure 6.6. The TWR, previously implemented, is performed between the mobile node and each of the anchor nodes. Having more anchor nodes, better position estimate can be achieved. A separate host computer can be used for the mobile node. A separate WLAN link was used for data communication between the anchor nodes and the mobile node.

The logical architecture of the localization system is shown in Figure 6.7. The radios are connected to ranging and reflector servers. The anchor nodes transmit a reply on each frame received from the mobile node and are connected to the reflector server. On the other hand, the mobile node is connected to the ranging server, as shown in Figure 6.7.

The TWR is performed between the mobile node and the available anchor nodes in a round-robin fashion. The estimated distances to the anchor nodes are used for position estimation using trilateration.

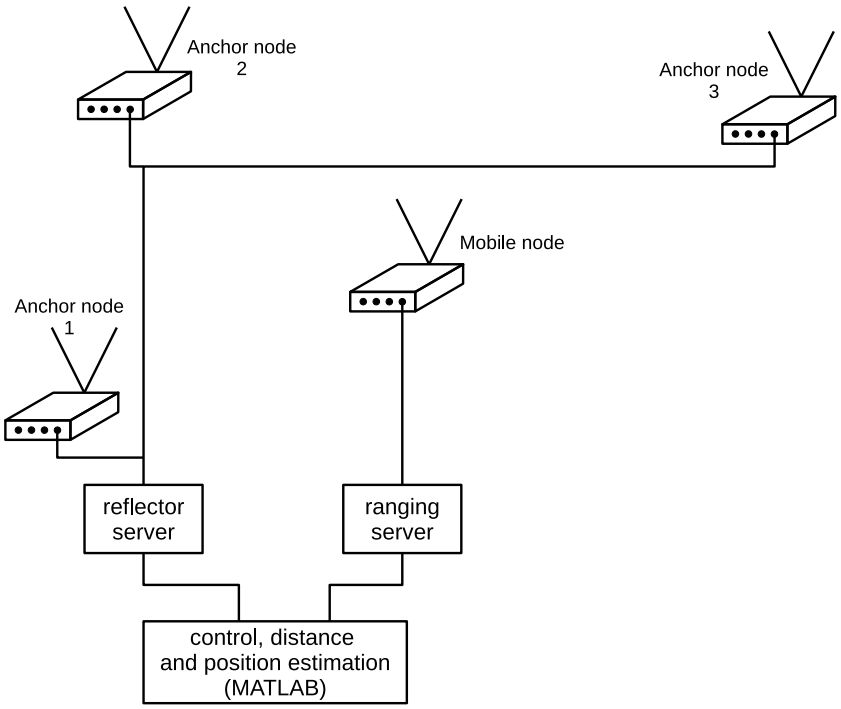


Fig. 6.7 Logical architecture of the developed positioning system

The localization process

The main localization process is controlled using a MATLAB script. The flowchart of the MATLAB script is shown in Figure 6.8. The script uploads the ranging frame into the mobile node and starts the ranging process. Also, parameters and threshold levels for preamble detection are uploaded in the mobile and anchor nodes. The transmitter power and receiver gain are set up to default values. The script issues a command and the TWR is performed by the mobile node and one of the anchor nodes. The script examines whether the TWR was successful. If a frame is not received back to the mobile node within a given timeout, the transmit power as well as the receive gain are increased slightly for the next TWR process. If a frame is received, the received samples and the coarse round trip time are transferred to MATLAB. These samples are used for estimation of the distance between the mobile node and the anchor nodes.

Graphical user interface

For the purpose of presentation of the ranging and localization results a GUI interface was developed in MATLAB. The developed GUI is shown on Figure 6.9. The estimated distances between the mobile node and the anchor nodes are shown on the right side, labeled as *R1*, *R2*

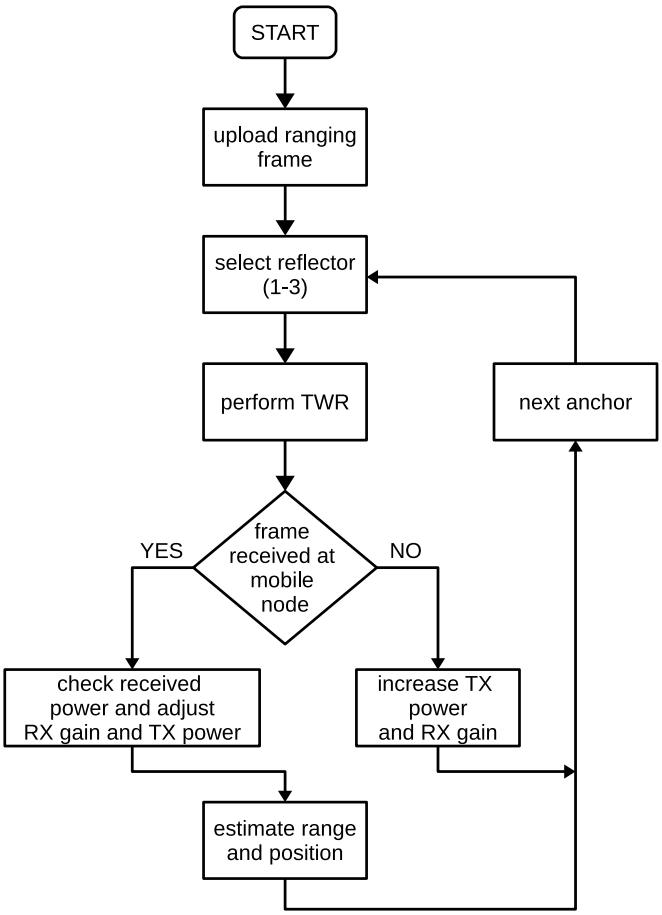


Fig. 6.8 Simplified flowchart of the MATLAB script used for performing the localization process

and *R3* in Figure 6.9. The transmit power gain as well as the receive gain are shown under *G1*, *G2* and *G3*. The relative value of the estimated received power is shown under the parameters *PWR1*, *PWR2* and *PWR3*. This relative power level is used to estimate the required transmit power and receive gain.

In lower right corner, different parameters can be changed. They include, threshold levels for the power and receive gain, threshold level for cross-correlation peak detection, reference position, etc.

Finally, different plots are shown on the GUI. They are primarily intended for debugging purposes. The first plot, *abs sig*, (left to right) presents the magnitude of the samples of the received frame. The second plot (*sig*) shows the in-phase and the quadrature signal samples of the frame. Further, the third plot, *crx*, is the cross-correlation function of the received frame with the locally generated ranging sequence. The cross-correlation peak can be clearly noticed in these plots. Finally, the last plot, *dist*, is a plot of the previous distance estimates. As can be noticed, these plots are shown for each anchor node.

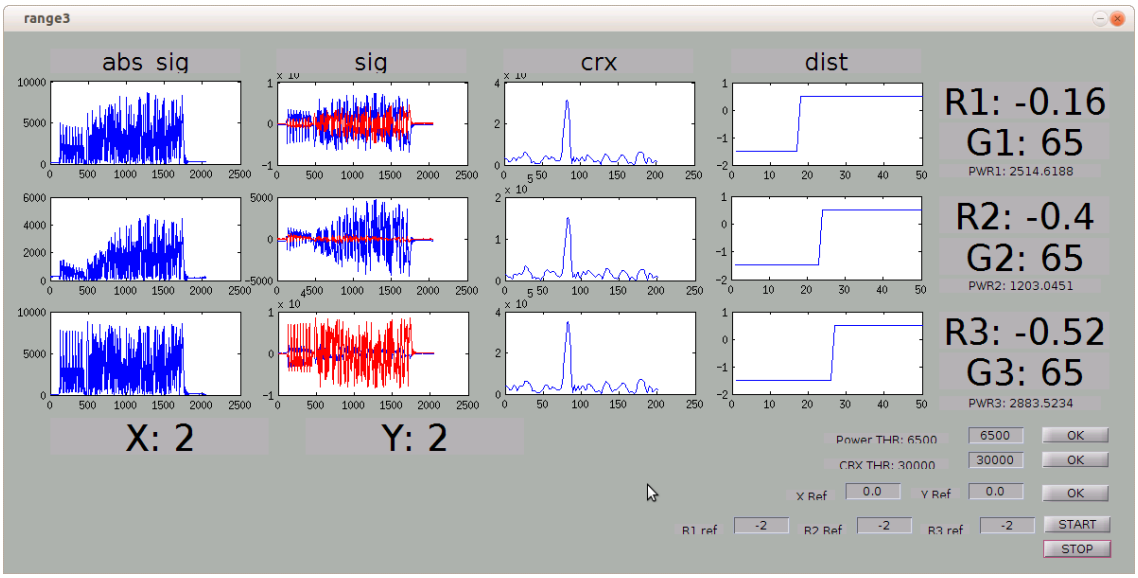


Fig. 6.9 MATLAB GUI used for visualizing the ranging and localization results

6.3 System setup, testing and results

The system is tested in ranging as well as in positioning scenarios. The estimated distances are compared to the reference distance obtained by other means. The positioning functionality was tested, but the precision evaluation was not performed.

As discussed previously, for ranging purposes a frame shown in Figure 6.2 is used. The preamble consists of four copies of the short training symbols (STS) from the IEEE802.11a standard [1]. These STSs are repeated 8 times and interpolated by a factor of 2. A m-sequence with length of 1023 is used as a ranging sequence. It is BPSK modulated, interpolated by a factor of 2 and filtered using a SRRC filter with a roll-off factor of 1. The 3 dB bandwidth of the transmitted signal is 25 MHz. The sample rate is 100 MSps of complex samples. The reflecting node does not perform any filtering. Additional SRRC filtering is performed when the frame is received back at the initiating node.

To test the precision of the distance estimates, two SDR platforms are used. The first is the originating node and the second is the reflecting node. Both nodes are connected to the same host computer. The originating node is static and the reflecting node is moved further away from it. The tests are performed in a narrow hallway, with a width of approx. 2 meters and a height of approx. 4 meters. The length of the hallway is approx. 20 meters. The walls are made of bricks, with steel supports. There are wooden and metal doors from both sides of the hallway. The initiating node is placed in the middle of the hallway, about 2 meters from the end. The true distance is measured using a measurement tape. The ranging precision is tested

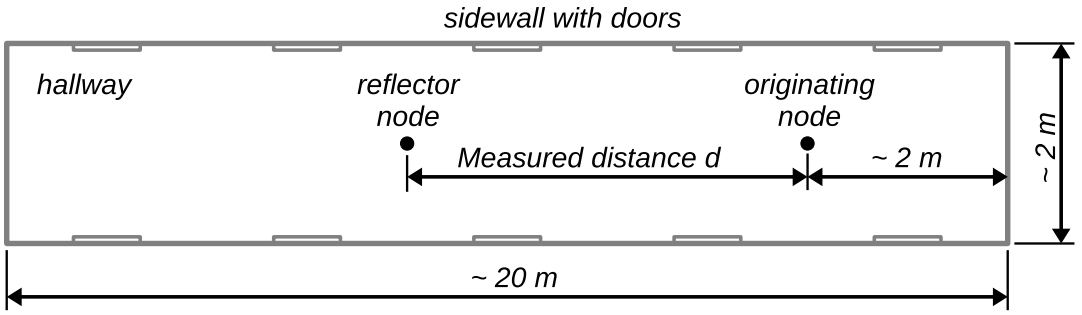


Fig. 6.10 Floor-plan of the hallway used for testing the ranging performances of the developed system

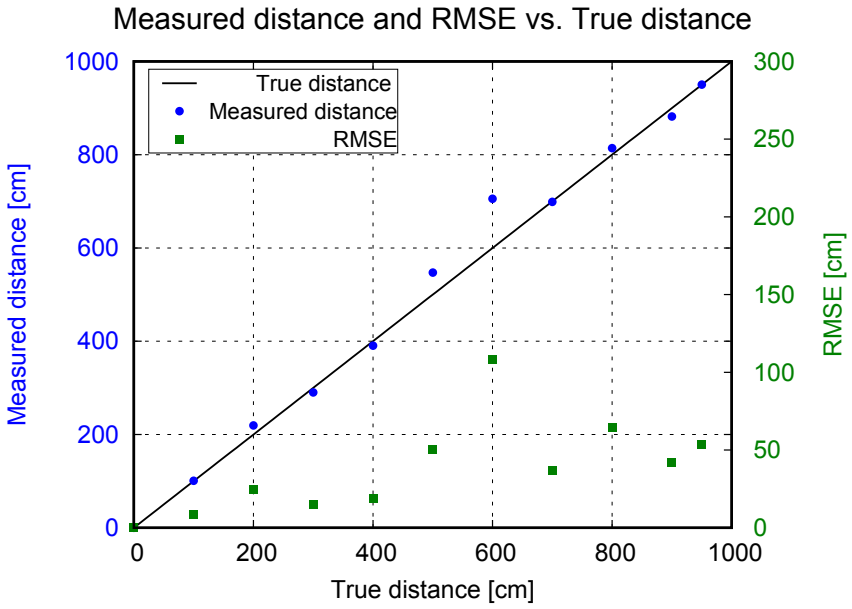


Fig. 6.11 Measured distance and distance RMSE as a function of true distance

for distances up to 9.5 meters, due to cable length limitations. The ranging scenario is shown in the Figure 6.10.

After the ranging frame is prepared and uploaded to the radio, the ranging process is started. The nodes are positioned as shown in Figure 6.10. The distance between the nodes is changed from 1 to 9 meters, with 1 meter step size. The last measurement is made at 9.5 meters. For each position, a large number of measurements (> 200) are made, in order to estimate the variance of the measurements.

Finally in Figure 6.11, the results from the measurements are summarized. The estimated distances, follow the true distances relatively well. For the distance of 6 meters, the error is slightly larger, probably due to additional reflections from the walls or the ground. Anyway, the root mean square error is 1 meter approximately in the worst case.

Table 6.3 Comparison with the other ranging and localization systems in the 2.4/5 GHz band.

System	Band	Channel BW [MHz]	Method	Precision (RMSE) [m]
This SDR system	2.4/5 GHz	25	TWR	<1
Schaffer et al. [98]	2.4 GHz	60	RADAR	0.33
Chronos [113]	2.4 GHz	20	TWR	0.1 (3 sec/meas.)
Jekabsons et al. [54]	2.4/5 GHz	20	RSS	3
Rea et al. [92]	2.4 GHz	20	ToA	3

In Table 6.3, the precision of this system is compared to the precision of other known and published systems. As can be noticed the two systems achieving higher precision either use larger channel bandwidth [98] or larger measurement time [113], i.e. 3 seconds per measurement.

6.4 Combined ranging and data communication operating in the 60 GHz band

Simultaneous data communication and ranging, or localization, is a key feature for future wireless systems. In this section a proposed approach for integration of ranging and data transmission functionality is presented.

An implementation of high data rate transmission and ranging on a common system, transmitting in the 60 GHz ISM band, is performed in order to verify the proposed approach. Standalone communication and ranging IP cores for this system are already available. Namely, the data transmission core, described in [24], was available for use, as well as the ranging core described in [83].

The data transmission and ranging cores support only point to point communication, as shown in Figure 6.12 and multi-user scenarios are not supported at all.

6.4.1 High data rate transmission core

The data transmission core for the 60 GHz band consists of a baseband processor and a MAC controller. In Figure 6.13, a basic block diagram of a complete data transceiver including this core is shown. The data transceivers are connected to host computers using the Gbit Ethernet port available on the MAC controller, as shown in Figure 6.12. This interface is used to transfer data from and to the host computer, as well as low data rate control information. The control information is used for modification of the modem parameters. The MAC controller

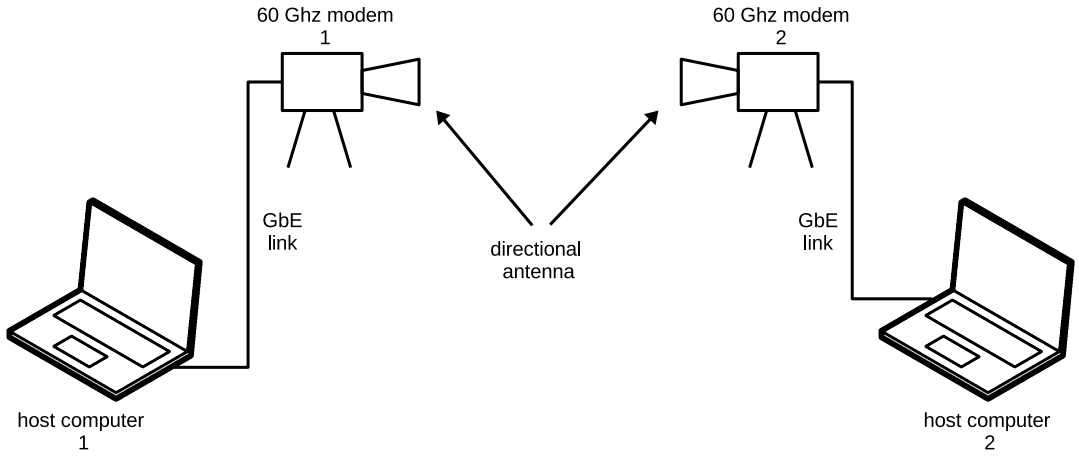


Fig. 6.12 Point-to-point communication between two 60 GHz nodes

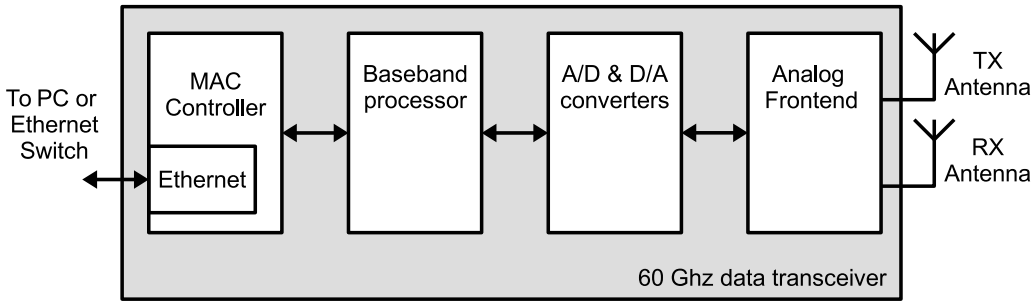


Fig. 6.13 Simplified block diagram of the 60 GHz transceiver

is connected to the baseband controller, which modulates the data to be transmitted and demodulates the received signal. The A/D and D/A converters are needed since the baseband processor has digital interfaces and the analog front-end is completely analog. The analog front-end translates the baseband signal in the 60 GHz band.

The main properties of the baseband core are summarized in the Table 6.4 as described in [24]. However this baseband core was not intended for ranging and localization and therefore, does not support any functionalities other than data transmission. The MAC controller, also does not support ranging and localization. Therefore, changes in the MAC core had to be introduced.

6.4.2 High precision localization core

The high precision ranging core was previously developed and implemented on the same 60 GHz system. The details for this core can be found in [83]. The core uses the so called Round

Table 6.4 Parameters of the baseband core

Parameter	Value
Channel bandwidth	2160 MHz
FFT bandwidth	2160 MHz
FFT size	1024
Sub-carrier spacing	2.1 MHz
Guard interval	119 ns
FFT period	474 ns
OFDM symbol time	593 ns
# data sub-carriers	768
# of pilot/zero sub-carriers	60/5
Nominal channel bandwidth	1757.11 MHz
Supported modulation schemes	BPSK; QPSK
Supported data rate	1.3 Gbps

Trip Phase method for range estimation. This is a Round Trip Time of Flight (RTToF) method, which uses the phase information from the received frame in order to estimate the ToF.

This method uses an orthogonal frequency division multiplexing (OFDM) frame, which is 256 symbols long. It is obtained by performing an inverse FFT on a frame defined in the frequency domain and consists 256 subcarriers. The frame is transmitted to the reflecting node and received back at the initiating node, as in a classical TWR method. The distance is estimated using the phase information from each sub-carrier. In order to increase the range resolution, an interpolation method called Zoom FFT [56] is performed. Zeros are added to the received waveform in the time domain and an FFT with additional points is performed. The result is a higher resolution FFT of the received waveform.

The round trip phase implementation works in a standalone mode, where one of the 60 GHz nodes is initiating and the other is a reflecting node. The initiating node starts the ranging process and after receiving the response from the reflecting node estimates the distance. In order to start the ranging process, the node must receive a command from the host computer which also configures the ranging parameters as well as the node role.

6.4.3 Proposed approach for integration of ranging and data transmission

A few different approaches for integration of ranging and data transmission functionality in a single system, are available [14, 18, 47]. Our main goal is to integrate both functionalities in a system which completely separates both cores and minimizes the necessary interaction

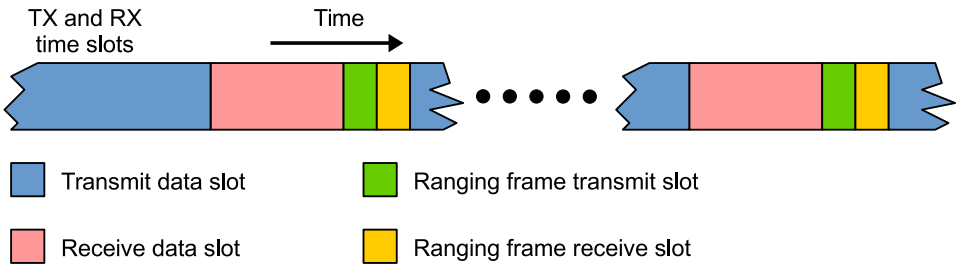


Fig. 6.14 Time division multiplexing of ranging and data communication.

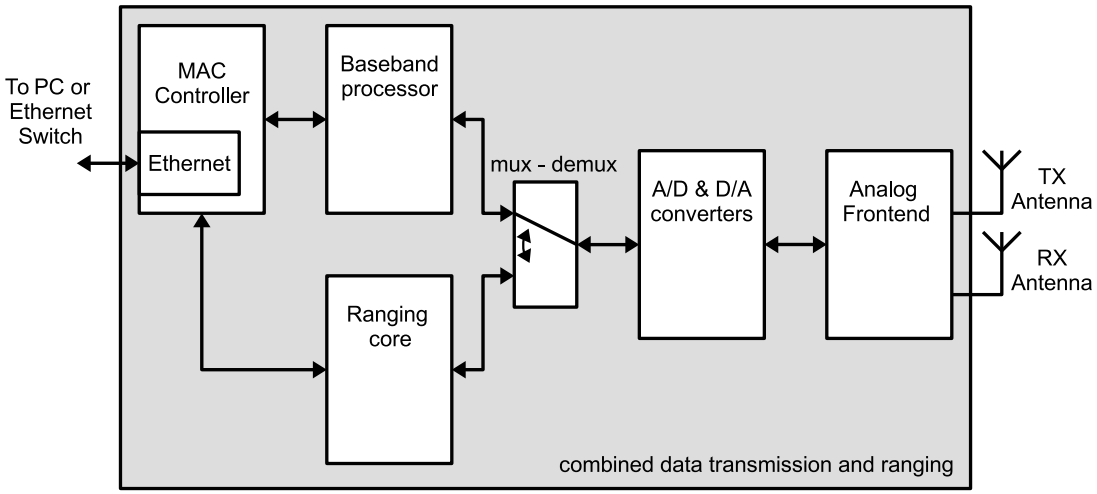


Fig. 6.15 The proposed integration approach

between them. The objective is to have an universal system in which different ranging and data transmission cores can be easily integrated.

We use Time Division Multiplexing (TDM) to integrate ranging and data transmission in a single system. The TDM uses separate time slots for data transmission and for ranging, as shown in Figure 6.14. Using this approach, the data transmission core and the ranging core functions do not intertwine, making the integration simpler. On the other hand, some of the functions must be duplicated, producing a not optimal system. This is not a significant drawback, since the main objective is to demonstrate joint ranging and high speed data communication with minimal implementation complexity involved. For product development, further optimization and granular resource sharing can be considered. The block diagram of the complete system is depicted in Figure 6.15. No interconnections between the baseband processor and the ranging core exist. The MAC controller controls the baseband processor as well as the ranging core. An additional multiplexer and demultiplexer (“mux - demux”), which connects either the baseband processor or the ranging core to the data converters is added. The control of both cores is performed using the MAC controller which is connected to the host computer. It schedules

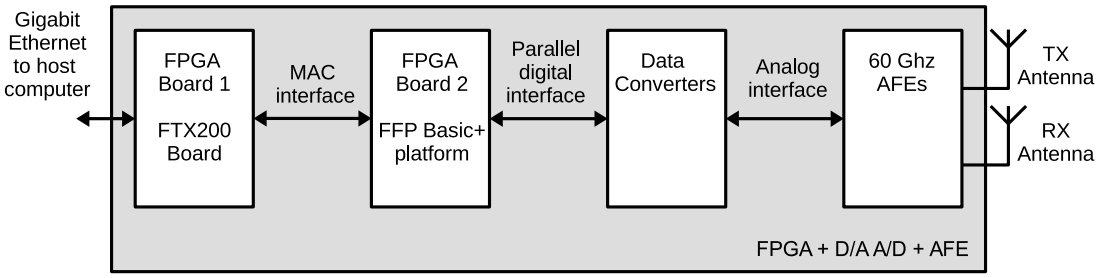


Fig. 6.16 Block diagram of the FPGA system used for integration

the data communication as well as the ranging time slots, depending on the user requirements. The existing MAC controller must be additionally modified in order to support ranging. It also should negotiate with the other wireless node in order to synchronize the ranging and data transmission slots. The ranging data is transferred to the host computer using the existing Gigabit Ethernet connection.

6.4.4 Integration of the precise ranging and high data rate transmission system

The integration of both functionalities, data transmission and ranging, is performed on FPGA devices. The data transmission and ranging cores are previously implemented as standalone systems on the same FPGA devices, which additionally simplifies the integration.

The system for integration consists of two FPGA boards, data converters (A/D and D/A) and 60 GHz analog front-ends. The first FPGA board uses Virtex 5 - XC5VFX200T FPGA[4] and runs the MAC controller. The second FPGA Board, is the FFP Basic+ board[5], which consists of four Virtex 5 devices, used for implementation of the baseband processor and the ranging functionality. A simplified block diagram of the overall system is shown in Figure 6.16. The data converters (A/D and D/A) are connected to this board and their digital interfaces are connected with two (out of the four available) FPGAs.

Integration of the systems on FPGA platform

The design of the system supporting ranging and high data rate transmission is split among four different FPGAs. The block diagram of the integrated system is shown in Figure 6.17. In this work, the *integration blocks* are developed and the other blocks were already available. The integration blocks are developed in order to enable the integration of the two cores, i.e. the

[4]The FPGA board is designed and manufactured by HighTechGlobal - http://www.hitechglobal.com/
[5]The FPGA board is designed and manufactured by IAF - http://www.iaf-bs.de/

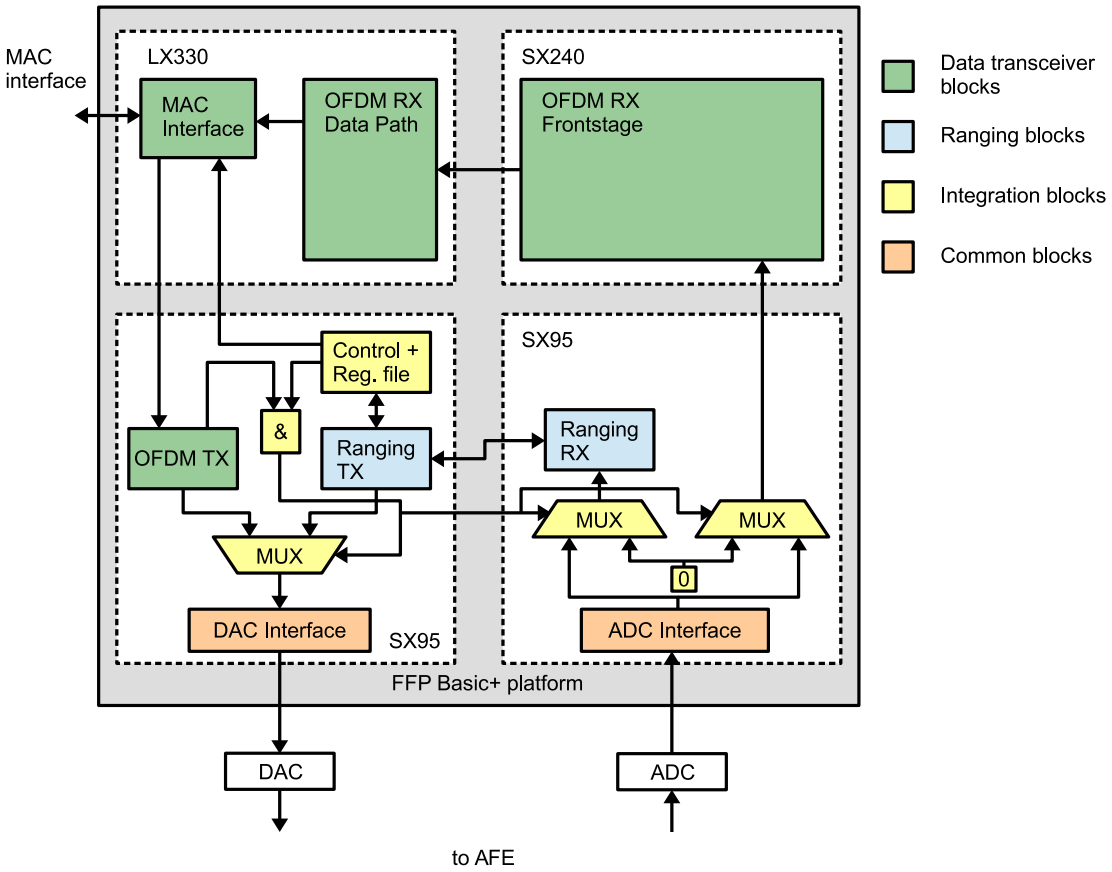


Fig. 6.17 Block diagram of the integrated ranging and baseband processing core

data transceiver and the ranging core. The MAC is implemented on a separate FPGA board and it is responsible for scheduling time slots for ranging and data transmission. The MAC communicates with the distant node in order to synchronize the slots. The synchronization error is less than 1 μs. The ranging functionality of the MAC controller [25] is not developed in this work.

A register file is also implemented in order to exchange data with the MAC layer and the host computer. The ranging parameters, as well as the commands are transferred using this register file. The control block controls the ranging process and communicates with the MAC controller. It starts the ranging process when a command is received and informs the MAC controller when the process is finished. It also controls the multiplexers (*MUX*). If no reply from the other node is received, i.e. a timeout expires, the control block reports an error to the MAC controller.

The ranging algorithm

The control block executes the ranging functionality of the implemented system. In order to perform ranging the MAC controller schedules a ranging time slot in which the TWR should be performed. Both nodes communicate and agree on the time slot start and duration. The MAC controller writes in the register file the maximum duration of the ranging time slot and issues a start command.

The control block waits in an idle state. When the start command is issued from the MAC controller, the ranging controller starts the ranging procedure. First the multiplexers connect the data converters to the ranging core. A timer expiring after the maximum allowed duration of the ranging time slot is started. The expiration of this timer automatically connects the baseband processor to the data converters.

The ranging controller sets up the ranging parameters and waits for a short guard interval. A guard interval is needed since parts of a data frame from the previous data transmission slot, can fall into the ranging slot, due to multipath propagation. After expiration of the guard interval, the ranging core transmits the ranging sequence to the other station. The distant station has a synchronizer, which detects the incoming ranging sequence (frame). The ranging core starts storing this sequence in RAM. When the required number of samples are acquired, they are retransmitted back to the initiating node.

No reconstruction or symbol detection is performed on the received waveform. The signal is not reconstructed, in order to avoid carrier frequency offset estimation and compensation in both nodes.

When the samples are received back at the initiating node, cross-correlation in frequency domain is performed and the cross-correlation peaks are identified. The round trip time is estimated. The reply time of the distant station has a fixed duration. A small error in the reply time would arise due to the clock inaccuracy at the nodes. Nevertheless, due to short reply time (<50 μs), this error would be negligible. The estimated distance value is stored in the register file and it is indicated to the MAC controller that the range estimation is finished and the result is available. In case that no received sequence is detected, in the allocated time slot, an error is reported to the MAC controller, after expiration of the reserved timeout.

In Figure 6.18, a time diagram of this scenario is shown. The two ranging windows scheduled by the MAC controller are not ideally aligned, and there is a “sync error”, which, as mentioned, is less that 1 μs. The introduced guard interval ensures that the transmissions would fall in the ranging time slots.

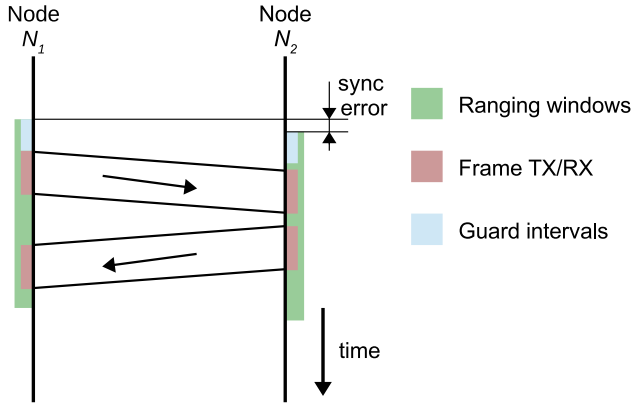


Fig. 6.18 Time diagram of the TWR implemented on the 60 GHz system

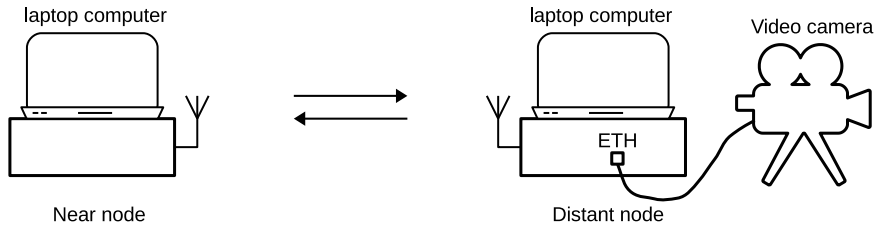


Fig. 6.19 The configuration of the test scenario.

6.5 Testing of the 60 GHz ranging and data transmission system

The developed high data rate transmission and ranging system was tested using two nodes. The configuration is shown in Figure 6.19. An industrial video camera is used as a high data rate source. The camera, can supply different data rates, by changing the video capture parameters. The host computers connected to the nodes monitor the parameters of the system as well as the estimated distance and the achieved data rate.

Testing of the system was performed in a room with dimensions of $10 \times 6 \times 3.5$ meters. The room has brick walls on three sides and windows at the fourth side. The floor and ceiling are made of concrete with steel reinforcements. The reference distances were measured and marked on the floor. The separation between the nodes is changed in steps of 20 centimeters. Both nodes are placed on carts in order to move them freely. The antennas are placed at a height of approx. 85 *cm* and are aligned. The antennas have a beamwidth of approx. 30 degrees.

A few hundred distance measurements are performed for each position between the nodes. Mean value and standard deviation of the distance is calculated for each position of the nodes.

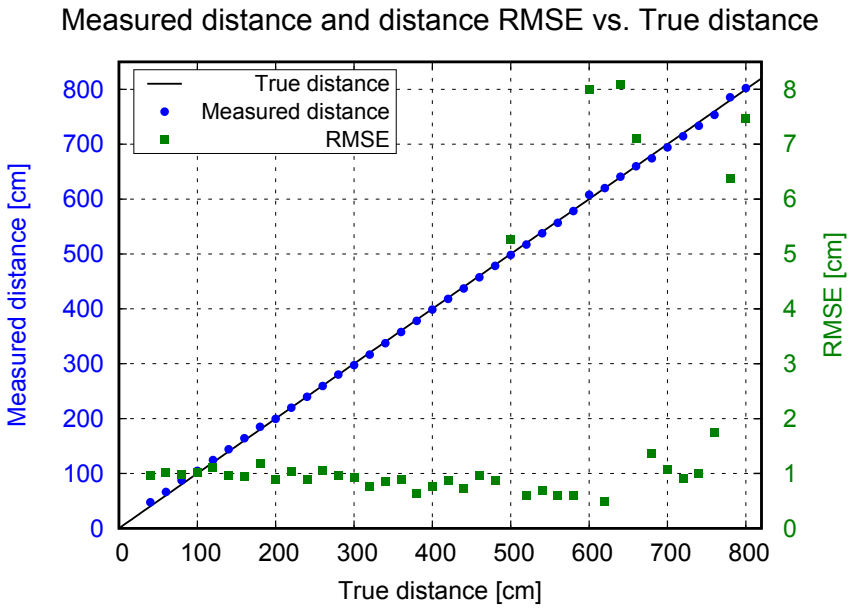


Fig. 6.20 Measured distance and distance RMSE as a function of true distance

During the measurements, a video stream with approx. 1 Gbps was transmitted from one node to the other. About 20 range measurements were performed per second. The duration of each ranging measurement was less than 80 μs. This duration includes the time needed for estimation of the distance. This basically leaves 99.8 % of the time, available for data communication.

The obtained results are shown in Figure 6.20. The measured distance follows the true distance quite accurately. The TWR accuracy and precision do not depend on the distance as shown in the presented results. This is possible only if the SNR is above a certain limit.

The standard deviation of the measurement error is mainly below 1 cm. At 5 meters and above large ranging errors can be observed. The source of this error is the ground reflection which becomes significant at larger distances.

This example shows that the functionality of both systems is not affected by their integration and the performance is not notably reduced. The loss of data rate is negligible, i.e. 0.2 %, which is insignificant from a user perspective. Regarding the ranging performance, the precision is not even slightly affected.

6.5.1 Comparison with similar wideband systems

In this work, we compare the ranging and data communication capabilities of our system, with the capabilities of other similar systems. The comparison is also performed with the

Table 6.5 Comparison with other ranging and localization systems in the 60 GHz and UWB band.

System	Band	Channel BW [MHz]	Method	Precision [cm]
This work 60 GHz system	60 GHz	2000	TWR	1.5
Fang et al. [29]	60 GHz	1500	RSS	3
Bocquet et al. [11]	60 GHz	3000	TDoA	12
Indirayanti et al. [53]	60 GHz	6000	ToA	0.2
Malajner et al. [76]	UWB	500	TWR	10
Kreiser et al. [65]	UWB	500	TWR	4

UWB systems, since they also use large channel bandwidth and have ranging and localization capabilities similar to the 60 GHz systems.

Not many 60 GHz systems support simultaneous localization and ranging or positioning. They are not yet widely used due to the high cost and low availability. Nevertheless, it is expected that with the upcoming developments the costs associated with this technology are going to be significantly reduced making it interesting for future research.

In Table 6.5, the main characteristic of the developed 60 GHz system, compared to other similar systems is shown. As can be noticed, the ranging precision is strongly dependent on the channel bandwidth used. The systems working in the 60 GHz band utilize larger channel bandwidth and, therefore, achieve higher ranging precision. The UWB systems, on the other hand, utilize smaller channel bandwidth of 500 MHz. They are also power limited, since the UWB channels overlap with other narrowband channels. These are the two main reasons for lower ranging precision of the UWB systems. Nevertheless, the IR UWB systems are extremely simple and cheap which makes them attractive for use in ranging and localization applications.

Chapter 7

Conclusion and future work

This work focuses on methods for indoor radio frequency ranging and localization. They are actively being investigated in the last decade in order to develop localization systems intended to cover areas in which global navigation satellite systems are not available or offer poor performances. These methods are similar to the methods used in earlier navigation systems or radar. Anyway, they must offer higher precision and accuracy since they can cover only small areas. These methods must also have lower complexity since they should be implemented in low cost devices deployed indoors. Integration of localization functionalities in wireless data transceivers, e.g. WiFi, would additionally reduce the cost of the indoor localization systems.

The main focus is towards solutions that would not additionally increase the hardware complexity or the cost of the ranging or localization system. A localization system for the 2.4/5 GHz is proposed and implemented. It uses a software defined radio platform and requires minimal resources. Further, an approach for integration of a data transmission and ranging in a single system is also proposed and demonstrated.

In this work, two approaches for increasing precision and accuracy of indoor ranging and localization systems are proposed. Firstly, the Modified Equivalent Time Sampling approach is proposed to reduce the ranging error resulting from the finite sample rate of the used data converters. A similar approach is used in the GPS system, but it is not easily applicable in systems supporting high data rate transmission. The proposed approach oversamples the received waveform using signal processing techniques and can be used in every system that supports data communication. It involves oversampling of the waveform of interest, e.g. only those used for ranging purposes, and allows easy selection of the oversampling factor depending on the needed precision and accuracy. The approach increases the precision and accuracy at the cost of additional measurement time. The choice of the oversampling factor determines the trade-off.

The second approach presented in this work is used in cooperative localization scenarios. In these scenarios, if the N-Way (or Multi-Way) ranging approach is used, the ranging error due to clock frequency offset between the nodes can become extremely large. The approach proposed in this work, uses the information about the clock frequency offset available at each node to compensate the ranging errors. It has been shown that this error can be significantly reduced without introducing additional complexity in the wireless nodes. The number of required transmissions is lower compared to other similar methods.

Finally, an implementation of ranging and localization in the 2.4/5 GHz band, as well as integration of ranging and high speed data communication in the 60 GHz band has been performed. The ranging and localization in the 2.4/5 GHz band was achieved using commercial software defined radio platforms. Parts of the FPGA configuration were changed in order to fulfill the precise timing requirements for TWR. The localization is implemented using trilateration with three anchor nodes. The ranging accuracy achieved is about 1 meter which is an acceptable value for indoor localization applications.

An approach for integration of a ranging and data transmission functionality in a single system is proposed and implemented. The implementation is performed on a system transmitting in the 60 GHz band since data transmission and ranging cores were already available for it. The proposed approach allows seamless integration of standalone wireless data transmission and ranging cores in a common system. The interaction between the cores is minimal and the needed changes of the cores are negligible. The performances and the functionalities of the available cores in the joint system are not even slightly affected.

The main contribution of this work is the development of methods for improvement of the ranging precision in indoor environments. The RF ranging and localization methods are generally known and are commonly used for outdoor application. Anyway, they cannot be directly used in indoor applications, since, the ranging errors would be to large. The methods developed in this work, significantly contribute towards achieving this general objective. Another contribution is the adaptation and simplification of this methods, allowing their implementation on existing data transmission systems, with minimum additional effort and added complexity. This is an important objective, since the indoor localization systems should not be costly or extremely complex.

Future work

The future work is mainly focused on localization in multi user scenarios. The TWR methods are a good trade off between complexity and ranging precision. However, in multi user

localization scenarios, this method becomes quite ineffective. The main disadvantage is the large number of transmissions needed for localization purposes. In mobile scenarios, where high position update rates are needed, the TWR can lead to significant channel usage for localization purposes. Therefore, methods that reduce the number of required transmissions are of interest. Our future work is focused on ToA based methods since they can achieve high ranging and localization precision. These methods require strict synchronization of the anchor nodes and, therefore, the focus of the future work also involves different synchronization methods for wireless nodes.

References

[1] (1999). IEEE Standard for Telecommunications and Information Exchange Between Systems - LAN/MAN Specific Requirements - Part 11: Wireless Medium Access Control (MAC) and physical layer (PHY) specifications: High Speed Physical Layer in the 5 GHz band. *IEEE Std 802.11a-1999*, pages 1–102.

[2] Abhayawardhana, V., Wassell, I., Crosby, D., Sellars, M., and Brown, M. (2005). Comparison of empirical propagation path loss models for fixed wireless access systems. In *Vehicular Technology Conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, volume 1, pages 73–77 Vol. 1.

[3] Angelis, A. D., Dwivedi, S., and Händel, P. (2012). Development of a radio front end for a UWB ranging embedded test bed. In *2012 IEEE International Conference on Ultra-Wideband*, pages 31–35.

[4] Arora, S. and Barak, B. (2009). *Computational complexity: a modern approach*. Cambridge University Press.

[5] Barnes, J., Rizos, C., Wang, J., Small, D., Voigt, G., Gambale, N., et al. (2002). High precision indoor and outdoor positioning using LocataNet. *Positioning*, 1(05).

[6] Bellini, S. and Tartara, G. (1974). Bounds on Error in Signal Parameter Estimation. *Communications, IEEE Transactions on*, 22(3):340–342.

[7] Biglieri, E. (2005). *Coding for wireless channels*. Springer Science & Business Media.

[8] Bisio, I., Lavagetto, F., Marchese, M., and Sciarrone, A. (2013). Energy efficient WiFi-based fingerprinting for indoor positioning with smartphones. In *2013 IEEE Global Communications Conference (GLOBECOM)*, pages 4639–4643.

[9] Björck, Å. (1996). *Numerical methods for least squares problems*. SIAM.

[10] Blanchard, W. F. (1991). Air Navigation Systems Chapter 4. Hyperbolic Airborne Radio Navigation Aids – A Navigator's View of their History and Development. *Journal of Navigation*, 44:285–315.

[11] Bocquet, M., Loyez, C., Lethien, C., Deparis, N., Heddebaut, M., Rivenq, A., and Rolland, N. (2010). A multifunctional 60-GHz system for automotive applications with communication and positioning abilities based on time reversal. In *The 7th European Radar Conference*, pages 61–64.

[12] Brannon, B. (1998). Basics of Designing a Digital Radio Receiver (Radio 101). *Analog Devices*.

[13] Brethour, V. (2005). Two Way Ranging Using Tracking Information to manage crystal offsets. *Contribution 802.15-05-0336-00-004a to the IEEE 802.15.4a Ranging Subcomittee*.

[14] Burnsweig, J. and Wooldridge, J. (1973). Ranging and Data Transmission Using Digital Encoded FM-"Chirp" Surface Acoustic Wave Filters. *IEEE Transactions on Microwave Theory and Techniques*, 21(4):272–279.

[15] Chazan, D., Zakai, M., and Ziv, J. (1975). Improved Lower Bounds on Signal Parameter Estimation. *Information Theory, IEEE Transactions on*, 21(1):90–93.

[16] Chen, C., Lai, H. Q., Chen, Y., and Liu, K. J. R. (2015). Accurate carrier frequency offset estimation in time-reversal communications. In *2015 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, pages 200–204.

[17] Chen, C., Wu, S., Meng, S., Chen, J., Fang, G., and Yin, H. (2011). Application of equivalent-time sampling combined with real-time sampling in UWB through-wall imaging radar. In *Instrumentation, Measurement, Computer, Communication and Control, 2011 First International Conference on*, pages 721–724. IEEE.

[18] Ciurana, M., Barcelo-Arroyo, F., and Izquierdo, F. (2007). A ranging system with IEEE 802.11 data frames. In *2007 IEEE Radio and Wireless Symposium*, pages 133–136.

[19] Coulson, A., Williamson, A., and Vaughan, R. (1998). A statistical basis for lognormal shadowing effects in multipath fading channels. *Communications, IEEE Transactions on*, 46(4):494–502.

[20] Cramér, H. (1945). *Mathematical Methods of Statistics*. Princeton landmarks in mathematics and physics. Princeton University Press.

[21] Davis, M. (2011). Radar Frequencies and Waveforms.

[22] Di Porto, A., Guida, F., and Montolivo, E. (1992). Fast algorithm for finding primitive polynomials over GF(q). *Electronics Letters*, 28(2):118–120.

[23] Ding, H., Liu, W., Huang, X., and Zheng, L. (2013). TOA estimation in IR UWB ranging using rank statistics with energy detection receiver under harsh conditions. In *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1255–1260.

[24] Ehrig, M. and Petri, M. (2012). 60 GHz wireless broadband cable replacement for machine vision applications. In *2012 International Symposium on Signals, Systems, and Electronics (ISSSE)*, pages 1–6.

[25] Ehrig, M., Petri, M., Sark, V., Gutierrez Teran, J., and Grass, E. (2014). Combined high-resolution ranging and high data rate wireless communication system in the 60 GHz band. In *Positioning, Navigation and Communication (WPNC), 2014 11th Workshop on*, pages 1–6.

[26] Ernst, M. and Bestmann, A. (2013). Kostengünstige Indoor Navigation und Lokalisierung durch hochpräzise Entfernungsmessung. Technical report, Silpion SolutionsCamp.

[27] Exel, R. and Bigler, T. (2014). ToA ranging using subsample peak estimation and equalizer-based multipath reduction. In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 2964–2969.

[28] Fang, B. T. (1986). Trilateration and extension to Global Positioning System navigation. *Journal of Guidance Control Dynamics*, 9:715–717.

[29] Fang, H. R., Cao, G. P., Gharavol, E. A., Tom, K., and Mouthaan, K. (2010). 60 GHz short range planar RSS localization. In *2010 Asia-Pacific Microwave Conference*, pages 1396–1399.

[30] Farnett, E. C. and Stevens, G. H. (1990). Pulse compression radar. *Radar handbook*, 2:10–1.

[31] Fereidoony, F., Sebt, M. A., Chamaani, S., and Mirtaheri, S. A. (2016). Model-based super-resolution time-delay estimation with sample rate consideration. *IET Signal Processing*, 10(4):376–384.

[32] Foerster, J., Green, E., Somayazulu, S., Leeper, D., Labs, I. A., Labs, I. A., Corp, I., and Corp, I. (2002). Ultra-Wideband Technology for Short-or Medium-Range Wireless Communications. *Intel Technology Journal*, 2:2001.

[33] Friis, H. T. (1946). A Note on a Simple Transmission Formula. *Proceedings of the IRE*, 34(5):254–256.

[34] Fuschini, F., Vitucci, E. M., Barbiroli, M., Falciasecca, G., and Degli-Esposti, V. (2015). Ray tracing propagation modeling for future small-cell and indoor applications: A review of current techniques. *Radio Science*, 50(6):469–485. 2015RS005659.

[35] Fyhn, K., Duarte, M. F. . f. . N., and Jensen, S. H. f. N. (2015). Compressive Parameter Estimation for Sparse Translation-Invariant Signals Using Polar Interpolation. *IEEE Transactions on Signal Processing*, 63(4):870–881.

[36] Galov, A., Moschevikin, A., and Voronov, R. (2011). Combination of RSS localization and ToF ranging for increasing positioning accuracy indoors. In *2011 11th International Conference on ITS Telecommunications*, pages 299–304.

[37] Gardner, F. (1986). A BPSK/QPSK Timing-Error Detector for Sampled Receivers. *IEEE Transactions on Communications*, 34(5):423–429.

[38] Gold, R. (1967). Optimal binary sequences for spread spectrum multiplexing (Corresp.). *Information Theory, IEEE Transactions on*, 13(4):619–621.

[39] Golomb, S. W. et al. (1982). *Shift register sequences*. Aegean Park Press.

[40] Green, M. P. (2005). N-Way Time Transfer ("NWTT") Method for Cooperative Ranging. *Contribution 802.15-05-0482-00-004a to the IEEE 802.15.4a Ranging Subcomittee*.

[41] Grewal, M. S., Weill, L. R., and Andrews, A. P. (2007). *Global positioning systems, inertial navigation, and integration*. John Wiley & Sons.

[42] Gyles, C. (1987). High-speed precision equivalent time sampling A/D converter and method. US Patent 4,654,584.

[43] Hach, R. (2005). Symetric Doubel Sided - Two Way Ranging. *Contribution 802.15-05-0334-00-004a to the IEEE 802.15.4a Ranging Subcomittee*.

[44] Hansen, T. and Mullen, G. L. (1992). Supplement to "Primitive polynomials over finite fields". *Math. Comp.*, 59(200):S47–S50.

[45] Hashemi, H. (1993). The indoor radio propagation channel. *Proceedings of the IEEE*, 81(7):943–968.

[46] Heiskala, J. and Terry, P. J. (2001). *OFDM Wireless LANs: A Theoretical and Practical Guide*. Sams, Indianapolis, IN, USA.

[47] Heldwein, C., Jekel, R., Sampson, S., and Zupan, J. (1980). Ranging system and method for determining the range of a vehicle from a plurality of reference points. US Patent 4,229,737.

[48] Helleseth, T. and Kumar, P. (1999). Pseudonoise Sequences. In Suthersan, S. S., editor, *Mobile Communications Handbook*. CRC Press LLC.

[49] Houghton, R. L., Strelich, T. P., Cluff, C. M., and Valine, J. C. (1998). Performance assessment of GPS augmentation systems. In *OCEANS '98 Conference Proceedings*, volume 3, pages 1340–1343 vol.3.

[50] Humphrey, D. and Hedley, M. (2008). Super-Resolution Time of Arrival for Indoor Localization. In *2008 IEEE International Conference on Communications*, pages 3286–3290.

[51] Humphrey, D. and Hedley, M. (2009). Prior Models for Indoor Super-Resolution Time of Arrival Estimation. In *VTC Spring 2009 - IEEE 69th Vehicular Technology Conference*, pages 1–5.

[52] Ianniello, J. (1986). Large and small error performance limits for multipath time delay estimation. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(2):245–251.

[53] Indirayanti, P., Ayhan, T., Verhelst, M., Dehaene, W., and Reynaert, P. (2014). A 60GHz transmitter in 40nm CMOS achieving mm-precision for discrete-carrier localization. In *ESSCIRC 2014 - 40th European Solid State Circuits Conference (ESSCIRC)*, pages 291–294.

[54] Jekabsons, G., Kairish, V., and Zuravlyov, V. (2011). An analysis of Wi-Fi based indoor positioning accuracy. *Scientific Journal of Riga Technical University. Computer Sciences*, 44(1):131–137.

[55] Jing, L., Liang, P., Maoyong, C., and Nongliang, S. (2008). Super-resolution time of arrival estimation for indoor geolocation based on IEEE 802.11 a/g. In *2008 7th World Congress on Intelligent Control and Automation*, pages 6612–6615.

[56] Jinming, L. and Huaiqiao, Y. (1995). Zoom FFT Spectrum by Fourier Transform [J]. *Journal of Vibration Engineering*, 2.

[57] Kanjilal, P. and Engineers, I. o. E. (1995). *Adaptive Prediction and Predictive Control*. Control, Robotics and Sensors Series. P. Peregrinus.

[58] Kaplan, E. (1996). *Understanding GPS: Principles and Applications*. Artech House telecommunications library. Artech House.

[59] Karl, H. and Willig, A. (2005). *Protocols and Architectures for Wireless Sensor Networks*. John Wiley & Sons.

[60] Kasami, T. (1966). Weight distribution formula for some class of cyclic codes. *Coordinated Science Laboratory Report no. R-285*.

[61] Keener, R. W. (2006). Statistical theory: notes for a course in theoretical statistics.

[62] Kerl, J. (2004). Computation in finite fields. *Arizona State University and Avnet Design Services*, 3.

[63] Khanzada, T. J. S., Ali, A. R., and Omar, A. S. (2008). Time Difference of Arrival estimation using super resolution algorithms to minimize Distance Measurement Error for indoor positioning systems. In *2008 IEEE International Multitopic Conference*, pages 443–447.

[64] Kisel, A. V. (1999). An extension of pulse shaping filter theory. *IEEE transactions on communications*, 47(5):645–647.

[65] Kreiser, D., Martynenko, D., Klymenko, O., and Fischer, G. (2015). Simple and efficient localization method for IR-UWB systems based on two-way ranging. In *2015 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, pages 1–4.

[66] Kwak, M. and Chong, J. (2010). A new Double Two-Way Ranging algorithm for ranging system. In *2010 2nd IEEE InternationalConference on Network Infrastructure and Digital Content*, pages 470–473.

[67] Lanzisera, S. and Pister, K. (2008). Burst Mode Two-Way Ranging with Cramer-Rao Bound Noise Performance. In *Global Telecommunications Conference, 2008. IEEE GLOBECOM 2008. IEEE*, pages 1–5.

[68] Lehmann, E. and Casella, G. (2003). *Theory of Point Estimation*. Springer Texts in Statistics. Springer New York.

[69] Li, J. and Wu, R. (1998). An efficient algorithm for time delay estimation. *Signal Processing, IEEE Transactions on*, 46(8):2231–2235.

[70] Li, X. and Pahlavan, K. (2004). Super-resolution TOA estimation with diversity for indoor geolocation. *IEEE Transactions on Wireless Communications*, 3(1):224–234.

[71] Liu, N. (2010). *Performance of Time Delay Estimation and Range-Based Localization in Wireless Channels*. PhD thesis, UC Riverside.

[72] Liu, Q., Wang, Y., and Fathy, A. E. (2012). A compact integrated 100 GS/s sampling module for UWB see through wall radar with fast refresh rate for dynamic real time imaging. In *Radio and Wireless Symposium (RWS), 2012 IEEE*, pages 59–62. IEEE.

[73] Liu, X., Zhang, H., Gulliver, T. A., and w. Xu, L. (2015). Ranging performance with 60 GHz signals in indoor residential environments. In *2015 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing (PACRIM)*, pages 70–73.

[74] Lowrie, W. (2007). *Fundamentals of Geophysics*. Cambridge University Press.

[75] Lyons, R. G. (2004). *Understanding Digital Signal Processing (2Nd Edition)*. Prentice Hall PTR, Upper Saddle River, NJ, USA.

[76] Malajner, M., Planinšič, P., and Gleich, D. (2015). UWB ranging accuracy. In *2015 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pages 61–64.

[77] Manickam, T., Vaccaro, R., and Tufts, D. (1994). A least-squares algorithm for multipath time-delay estimation. *Signal Processing, IEEE Transactions on*, 42(11):3229–3233.

[78] McCluskey, P., O'Connor, C., and Nathan, K. (2001). Evaluating the Performance and Reliability of Embedded Computer Systems for Use in Industrial and Automotive Temperature Ranges. *CALCE Electronic Products and Systems Centre, University of Maryland*.

[79] Meel, J. (1999). Spread Spectrum (SS). Technical report, Vlaams Instituut voor de bevordering van het Wetenschappelijk Technologisch onderzoek in de industrie.

[80] Mueller, K. and Muller, M. (1976). Timing Recovery in Digital Synchronous Data Receivers. *IEEE Transactions on Communications*, 24(5):516–531.

[81] Murphy Jr., W. S. and Hereman, W. (1995). Determination of a Position in Three Dimensions using Trilateration and Approximate Distances. *Technical Report*, MCS-95-07.

[82] Obata, K., Mizutani, K., and Harada, H. (2016). Carrier frequency offset estimation scheme for IEEE 802.15.4g based wide area Wi-SUN systems. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6.

[83] Ohlemueller, T., Winkler, F., and Grass, E. (2010). Radio localization in OFDM networks using the round trip phase. In *Positioning Navigation and Communication (WPNC), 2010 7th Workshop on*, pages 23–27.

[84] Patwari, N. (2005). *Location estimation in sensor networks*. PhD thesis, Citeseer.

[85] Patwari, N., Ash, J. N., Kyperountas, S., Hero, A. O., Moses, R. L., and Correal, N. S. (2005). Locating the nodes: cooperative localization in wireless sensor networks. *IEEE Signal Processing Magazine*, 22(4):54–69.

[86] Pelka, M., Bollmeyer, C., and Hellbrück, H. (2015). Indoor localization based on bi-phase measurements for wireless sensor networks. In *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1362–1367.

[87] Pivato, P., Dalpez, S., and Macii, D. (2012). Performance evaluation of Chirp Spread Spectrum ranging for indoor embedded navigation systems. In *7th IEEE International Symposium on Industrial Embedded Systems (SIES'12)*, pages 307–310.

[88] Ranganathan, G., v. Boegel, G., Meyer, F., and Grabmaier, A. (2016). A Survey of UWB Technology Within RFID Systems and Wireless Sensor Networks. In *Smart SysTech 2016; European Conference on Smart Objects, Systems and Technologies*, pages 1–7.

[89] Rao, C. and Das Gupta, S. (1994). *Selected papers of C.R. Rao*. Number v. 1 in Selected Papers of C.R. Rao. Wiley.

[90] Rao, R. C. (1945). Information and the accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, 37:81–91.

[91] Rappaport, T. S. et al. (1996). *Wireless communications: principles and practice*, volume 2. prentice hall PTR New Jersey.

[92] Rea, M., Cordobés de la Calle, H., Giustiniano, D., and Lenders, V. (2016). Robust WiFi Time-of-Flight Positioning System.

[93] Rice, M. and Saquib, M. (2016). On frequency offset estimation errors in coherent detection over equalized ISI channels. In *MILCOM 2016 - 2016 IEEE Military Communications Conference*, pages 1292–1297.

[94] Rizos, C., Kealy, A., Li, B., Choudhury, M., Choy, S., and Feng, Y. (2014). Locata's VRay™ Antenna Technology – Multipath Mitigation for Indoor Positioning. In *Proceedings of the 27th International Technical Meeting of The Satellite Division of the Institute of Navigation (ION GNSS+ 2014)*, pages 1629–1635.

[95] Saleh, A. and Valenzuela, R. (1987). A Statistical Model for Indoor Multipath Propagation. *Selected Areas in Communications, IEEE Journal on*, 5(2):128–137.

[96] Sark, V. and Grass, E. (2015). Configurable Software Defined Radio System For Two Way Time Of Flight Ranging. In *1st URSI Atlantic Radio Science Conference (URSI AT-RASC)*.

[97] Sawada, H., Shoji, Y., and Choi, C.-S. (2006). Proposal of novel statistic channel model for millimeter wave WPAN. In *Microwave Conference, 2006. APMC 2006. Asia-Pacific*, pages 1855–1858.

[98] Schaffer, B., Kalverkamp, G., and Biebl, E. (2014). A 2.4 GHz high precision local positioning system based on cooperative roundtrip time of flight ranging. In *GeMiC 2014; German Microwave Conference*, pages 1–4.

[99] Seybold, J. S. (2005). *Introduction to RF propagation*. John Wiley & Sons.

[100] Shaked, R., Shlezinger, N., and Dabora, R. (2016). Carrier frequency offset estimation for linear channels with periodic characteristics. In *2016 IEEE 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5.

[101] Shin, F., Zhiwei, L., and Xing, L. J. (2007). Symmetric Multi-Way Ranging for Wireless Sensor Network. In *Personal, Indoor and Mobile Radio Communications, 2007. PIMRC 2007. IEEE 18th International Symposium on*, pages 1–5.

[102] Shmaliy, Y. (2006). *Continuous-Time Signals*. Signals and Communication Technology. Springer Netherlands.

[103] Shparlinski, I. (1993). Finding irreducible and primitive polynomials. *Applicable Algebra in Engineering, Communication and Computing*, 4(4):263–268.

[104] Sudalaiyandi, S., Hjortland, H. A., Vu, T. A., Næss, O., and Lande, T. S. (2012). Continuous-time high-precision IR-UWB ranging-system in 90 nm CMOS. In *2012 IEEE Asian Solid State Circuits Conference (A-SSCC)*, pages 349–352.

[105] Tetley, L. and Calcutt, D. (1986). *Electronic aids to navigation*. Edward Arnold.

[106] Toda, A. P., Romeu, J., Jofre, L., and Flaviis, F. D. (2014). Beamforming antennas for 60 GHz positioning systems. In *The 8th European Conference on Antennas and Propagation (EuCAP 2014)*, pages 630–633.

[107] Trefethen, L. N. and Bau III, D. (1997). *Numerical linear algebra*, volume 50. Siam.

[108] Tsui, J. B.-Y. (2004). *Basic GPS Concept*.

[109] Tufvesson, F., Edfors, O., and Faulkner, M. (1999). Time and frequency synchronization for OFDM using PN-sequence preambles. In *Vehicular Technology Conference, 1999. VTC 1999-Fall. IEEE VTS 50th*, volume 4, pages 2203–2207. IEEE.

[110] Turin, G. (1960). An introduction to matched filters. *IRE Transactions on Information Theory*, 6(3):311–329.

[111] Van Trees, H. (2004). *Detection, Estimation, and Modulation Theory*. Number pt. 1 in Detection, Estimation, and Modulation Theory. Wiley.

[112] Vaseghi, S. V. (2006). *Advanced Digital Signal Processing and Noise Reduction*. John Wiley & Sons.

[113] Vasisht, D., Kumar, S., and Katabi, D. (2016). Decimeter-Level Localization with a Single WiFi Access Point. In *13th USENIX Symposium on Networked Systems Design and Implementation (NSDI 16)*, pages 165–178, Santa Clara, CA. USENIX Association.

[114] Wang, S., Shao, J., Tu, S., and Dai, J. (2015). Carrier frequency offset estimation for DSSS signals in the wired test of beamforming networks. In *TENCON 2015 - 2015 IEEE Region 10 Conference*, pages 1–6.

[115] Weiss, A. and Weinstein, E. (1983). Fundamental limitations in passive time delay estimation–Part I: Narrow-band systems. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 31(2):472–486.

[116] Xilinx, I. (2003). *Using Block RAM in Spartan-3 FPGAs*.

[117] Zhang, T., Zhang, Q., Xu, H., and Zhang, H. (2013). A practical ranging method using IR-UWB signals. In *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, pages 1839–1844.

[118] Zhang, W. and Yu, F. (2017). Off-the-Grid Compressive Time Delay Estimation via Manifold-based Optimization. *IEEE Communications Letters*, PP(99):1–1.

[119] Zhou, H., Malipatil, A. V., and Huang, Y.-F. (2006). Maximum-likelihood carrier frequency offset estimation for OFDM systems in fading channels. In *Wireless Communications and Networking Conference, 2006. WCNC 2006. IEEE*, volume 3, pages 1461–1464. IEEE.

[120] Ziv, J. and Zakai, M. (1969). Some lower bounds on signal parameter estimation. *Information Theory, IEEE Transactions on*, 15(3):386–391.

Appendix A

Distance estimation error of the different ranging methods

In this appendix the distance estimation error for all the methods listed in Table 5.1 and used in multiple node localization scenarios, is derived. The scenario shown in Figure 2.1b is used. There are N nodes and all the distances between these nodes should be estimated. There are a total of $D = \frac{1}{2}N(N-1)$ pairs of nodes and distances.

A.1 Two way ranging

Two way ranging can be used to estimate the distance between each pair of nodes shown in Figure 2.1b. The distance estimation error is given in (5.5). Assuming that all of the nodes have a reply time of t_{reply} and the frequency offset between the node i and j is $\sigma_{(ij)f}$, the distance estimation error would be

$$\varepsilon_{(ij)TWR} = \frac{1}{2} t_{reply} \sigma_{(ij)f} \tag{A.1}$$

To compare this error with the distance estimation error using the other errors, this error can be additionally simplified as

$$\varepsilon_{TWR} \propto \frac{1}{2} t_{reply} \sigma_{(NN)f} \tag{A.2}$$

where $\sigma_{(NN)f}$ is the average value of the frequency offset.

A.2 Compensated two way ranging

The distance estimation error can be corrected if the frequency offset between the two nodes in TWR is known. Therefore, the nodes can estimate the frequency offset in order to correct the distance error. Nevertheless, the frequency offset estimation contains an error $\varepsilon_{(ij)f}$ and, therefore, complete correction of the distance estimation error would not be possible. The distance estimation error is derived by correcting (A.1) and is given by

$$\varepsilon_{(ij)TWR_F} = \frac{1}{2} t_{reply} \sigma_{(ij)f} \varepsilon_{(ij)f} \tag{A.3}$$

if all of the nodes have the same t_{reply}. In the more general case, this error would be

$$\varepsilon_{TWR_F} = \frac{1}{2} t_{reply} \sigma_{(NN)f} \varepsilon_{(NN)f} \tag{A.4}$$

where $\sigma_{(NN)f}$ is the mean standard deviation of the frequency offset between two nodes and $\varepsilon_{(NN)f}$ is mean frequency offset estimation error.

A.3 Symmetrical double sided - two way ranging

In the SDS-TWR case the TWR would be performed twice, one starting from node i and the second time starting from node j. The ranging errors obtained in the both cases would be

$$\begin{aligned} \varepsilon_{(ij)SDS} &= \frac{1}{2} t_{(ij)tof} \sigma_{(i)f} + \frac{1}{2} t_{reply} \sigma_{(ij)f} \\ \varepsilon_{(ji)SDS} &= \frac{1}{2} t_{(jitof} \sigma_{(j)f} - \frac{1}{2} t_{reply} \sigma_{(ji)f} \end{aligned} \tag{A.5}$$

under assumption that both nodes have same t_{reply} and frequency offset estimation error is small. Both the SDS-TWR measurements are averaged and therefore the error is also averaged. The final error would be

$$\varepsilon_{(ij)SDS} = \frac{1}{2} t_{(ij)tof} \left(\sigma_{(i)f} + \sigma_{(j)f} \right) \tag{A.6}$$

This error in more general case can be written as

$$\varepsilon_{SDS} \propto t_{tof} \sigma_f \tag{A.7}$$

where the t_{tof} is the ToF and the σ_f is the clock frequency inaccuracy of the nodes.

A.4 N-Way ranging

The distance estimation error in the N-Way ranging method is calculated for two different cases. The first case is for neighboring nodes. The distance estimation error in this case is the same as in the TWR case given in (A.1).

In the case of non-neighboring nodes, the distance estimation error is given in (5.10). Assuming that all of the nodes have the same t_{reply} and rearranging (5.10), the distance estimation error is given as

$$\varepsilon_{(ij)NWR} = t_{reply}\left(\sigma_{(i)f}(j-i-1) - \frac{1}{2}\sum_{k=i+1}^{j}(\sigma_{(k)f}+\sigma_{(k+1)f})\right) \tag{A.8}$$

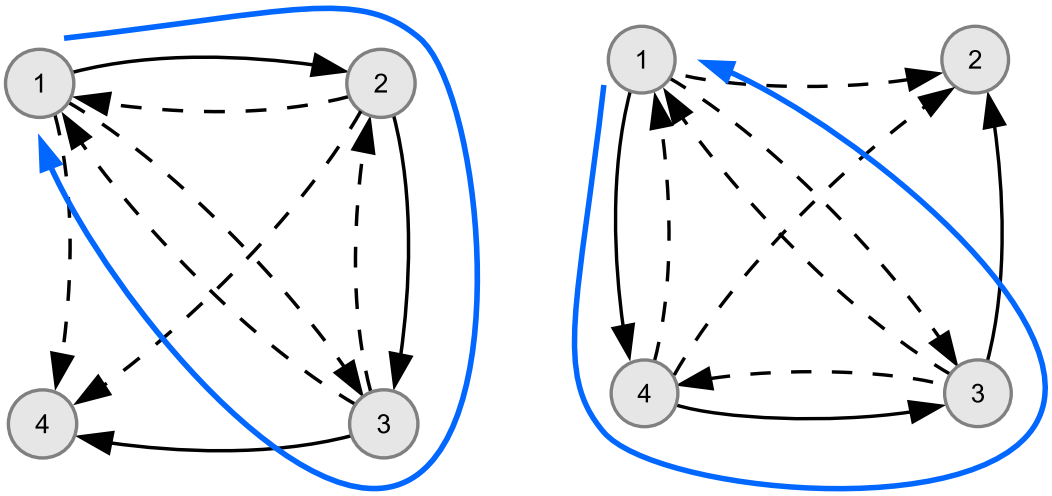
for distances between non neighboring nodes. For the neighboring nodes the variance of the distance estimates is equal to the TWR as in (A.1). Assuming that the frequency inaccuracies $\sigma_{(i)f}$ are not correlated, the distance estimation error in general case can be written as

$$\varepsilon_{NWR} \propto \frac{1}{2}Nt_{reply}\sigma_{(NN)f} \tag{A.9}$$

A.5 Symmetrical N-Way ranging

The symmetrical NWR as described in [101] would compensate only the ranging errors between neighboring nodes. The symmetrical NWR essentially performs a SDS-TWR between the neighboring nodes. Therefore, distance estimation error between them is the same as in (A.6). Nevertheless, the variance of the distance estimates between the non-neighboring nodes is not compensated. In Figure (A.1) the transmissions needed for estimation of the distance between the nodes 1 and 3 are shown. In the forward direction, Figure A.1a, the distance estimation error can be calculated using (5.7) and assuming that the ToF estimation error is fully compensated. In case where the t_{reply} of all of the nodes is equal, the distance estimation error between the nodes i and j is given as

$$\begin{aligned}\varepsilon_{ij} &= t_{(ij)reply}\left(\sum_{k=i+1}^{j}\sigma_i - \sum_{k=i+1}^{j}\sigma_k\right)\\ &= t_{(ij)reply}\left((j-i-1)\sigma_i - \sum_{k=i+1}^{j}\sigma_k\right)\end{aligned} \tag{A.10}$$



(a) NWR forward direction **(b)** NWR reverse direction

Fig. A.1 Estimation of the distance between the nodes 1 and 3 using symmetrical NWR

For the reverse direction, the the distance estimation error can be calculated similarly as

$$\begin{aligned}\varepsilon_{ij} = &= t_{(ij)reply}\left(\sum_{k=1}^{i-1}\sigma_i + \sum_{k=j}^{N}\sigma_i - \sum_{k=1}^{i-1}\sigma_k - \sum_{k=j}^{N}\sigma_k\right)\\ &= t_{(ij)reply}\left((N-(j-i)) - \sum_{k=1}^{i-1}\sigma_k - \sum_{k=j}^{N}\sigma_k\right)\end{aligned} \tag{A.11}$$

In the symmetric NWR, the measurement performed in the both directions are summed and averaged. Therefore, the errors in (A.10) and (A.10) are also summed and averaged. The final distance estimation error between nodes i and j is derived form (A.10) and (A.10) as

$$\varepsilon_{(ij)SNWR} = \frac{1}{2}t_{reply}\left((N+1)\sigma_{(i)f} - \sum_{k=1}^{N}\sigma_{(k)f} - \sigma_{(j)f}\right) \tag{A.12}$$

In the general case the distance estimation error would be

$$\varepsilon_{SNWR} \propto \frac{1}{2}Nt_{reply}\sigma_{(NN)f} \tag{A.13}$$

A.6 Proposed approach: compensated N-Way ranging

The proposed approach performs the classical NWR and uses the estimated frequency offset and performs correction of the estimated distance between the nodes. The distance estimation

error is not going to be fully compensated due to the frequency offset estimation error. In order to be able to easily the compare the errors in the different approaches, we assume that the frequency offset estimation error is ε_f for each pair of nodes. The distance estimation error between a pair of neighboring nodes would be the same as in (A.4). In case of a pair of non-neighboring nodes it would be the same as in (A.8) multiplied by ε_f as

$$\varepsilon_{(ij)CNWR} = \varepsilon_f t_{reply} \left(\sigma_{(i)f}(j-i-1) - \frac{1}{2} \sum_{k=i+1}^{j} (\sigma_{(k)f} + \sigma_{(k+1)f}) \right) \tag{A.14}$$

and generally

$$\varepsilon_{CNWR} \propto \frac{1}{2} \varepsilon_f N t_{reply} \sigma_{(NN)f} \tag{A.15}$$

Selbständigkeitserklärung

Ich erkläre, dass ich die Dissertation selbständig und nur unter Verwendung der von mir gemäß § 7 Abs. 3 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät, veröffentlicht im Amtlichen Mitteilungsblatt der Humboldt-Universität zu Berlin Nr. 126/2014 am 18.11.2014 angegebenen Hilfsmittel angefertigt habe.

Berlin, 28.02.2017
Vladica Sark