Algorithm-Architecture Co-Design for Digital Front-Ends in Mobile Receivers

Isael Diaz

LUND UNIVERSITY

Abstract

The methodology behind this work has been to use the concept of algorithm-hardware co-design to achieve efficient solutions related to the digital front-end in mobile receivers. It has been shown that, by looking at algorithms and hardware architectures together, more efficient solutions can be found; i.e., efficient with respect to some design measure. In this thesis the main focus have been placed on two such parameters; first reduced complexity algorithms to lower energy consumptions at limited performance degradation, secondly to handle the increasing number of wireless standards that preferably should run on the same hardware platform. To be able to perform this task it is crucial to understand both sides of the table, i.e., both algorithms and concepts for wireless communication as well as the implications arising on the hardware architecture.

It is easier to handle the high complexity by separating those disciplines in a way of layered abstraction. However, this representation is imperfect, since many interconnected "details" belonging to different layers are lost in the attempt of handling the complexity. This results in poor implementations and the design of mobile terminals is no exception. Wireless communication standards are often designed based on mathematical algorithms with theoretical boundaries, with few considerations to actual implementation constraints such as, energy consumption, silicon area, etc. This thesis does not try to remove the layer abstraction model, given its undeniable advantages, but rather uses those cross-layer "details" that went missing during the abstraction. This is done in three manners:

In the first part, the cross-layer optimization is carried out from the algorithm perspective. Important circuit design parameters, such as quantization are taken into consideration when designing the algorithm for Orthogonal Frequency Division Multiplexing (OFDM) symbol timing, Carrier Frequency Offset (CFO), and Signal-to-Noise Ratio (SNR) estimation with a

single bit, namely, the Sign-Bit. Proof-of-concept circuits were fabricated and showed high potential for low-end receivers. In the second part, the cross-layer optimization is accomplished from the opposite side, i.e., the hardware-architectural side. A Software Defined Radio (SDR) architecture is known for its flexibility and scalability over many applications. In this work a filtering application is mapped into software instructions in the SDR architecture in order to make filtering-specific modules redundant, and thus, save silicon area. In the third and last part, the optimization is done from an intermediate point within the algorithm-architecture spectrum. Here, a heterogeneous architecture with a combination of highly efficient and highly flexible modules is used to accomplish initial synchronization in at least two concurrent OFDM standards. A demonstrator was build capable of performing synchronization in any two standards, including LTE, WiFi, and DVB-H.

Any intelligent fool can make things bigger, more complex, and more violent. It takes a touch of genius — and a lot of courage to move in the opposite direction.

– Ernst Friedrich "Fritz" Schumacher (1911 − 1977)

Contents

Preface

This dissertation summarizes the work carried out during the period between August 2008 and April 2014, at the department of Electrical and Information Technology of Lund University. Most of the relevant results from these fruitful years have been published in scientific journals, international conferences, or publicly available research reports, which are listed as follows:

- DIAZ, I., WILHELMSSON, L., SOFOTASIOS, P., MIAO, Y., TAN, S., EDFORS, O., AND ÖWALL, V. "A new approach to Sign-Bit based parameter estimation in OFDM receivers," *submitted to Springer Circuits, Systems, and Signal Processing*, March 2014.

  ☞ The author parted from the original idea of Mr. Wilhelmsson, simulated various realistic scenarios, with the help of some co-authors developed a hardware architecture as proof of concept, while others contributed with important mathematical insights.

- DIAZ, I., ZHANG, C., HOLLEVOET, L., SVENSSON, J., RODRIGUES, J., WILHELMSSON, L., OLSSSON, T., VAN DER PERRE, L., AND ÖWALL, V. "A New Digital Front-End For Flexible Reception In Software Defined Radio", *Elsevier, Microprocessors and Microsystems: Embedded Hardware Design, invited to special issue on NORCHIP*, March 2014.

  ☞ The author was invited to submit an extended version of the work presented at NORCHIP in 2013. The additional material included supplementary implementation results and theoretical analysis.

- DIAZ, I., ZHANG, C., HOLLEVOET, L., SVENSSON, J., RODRIGUES, J., WILHELMSSON, L., OLSSSON, T., VAN DER PERRE, L., AND ÖWALL, V. "Next generation digital front-end for multi-standard concurrent reception", *NORCHIP Conference*, November, 2013.

  ☞ The author was the main integrator of the various parts conforming

the front-end presented in this article, the author was also the main responsible of the algorithm development of one crucial module in the design, i.e., the synchronization engine. Finally, the author was the main participant in the verification of the entire front-end.

- DIAZ, I., TORREA-DURAN, R., POLLIN, S., VAN DER PERRE, L., AND ÖWALL, V., "Selective Channelization on an SDR Platform for LTE-A Carrier Aggregation," *IEEE International Conference on Electronics, Circuits, and Systems*, October, 2012.

  ☞ The author carried out the study that resulted in the publication of this article with technical assistance and guidance from the co-authors.

- WILHELMSSON, L., DIAZ, I., OLSSON, T., AND ÖWALL, V., "Analysis of a novel low complex SNR estimation technique for OFDM systems," *IEEE Wireless Communications and Networking Conference*, March, 2011.

  ☞ The author performed performance simulations and hardware implication analysis.

- WILHELMSSON, L., DIAZ, I., OLSSON, T., AND ÖWALL, V., "Performance analysis of sign-based pre-FFT synchronization in OFDM systems," *IEEE 71st Vehicular Technology Conference*, May, 2010.

  ☞ The author provided with important aspects regarding practical implementation of the presented algorithms.

- DIAZ, I., WILHELMSSON, L., RODRIGUES, J., LÖFGREN, J., OLSSON, T., AND ÖWALL, V., "A sign-bit auto-correlation architecture for fractional frequency offset estimation in OFDM," *IEEE International Symposium on Circuits and Systems*, May, 2010.

  ☞ The author created a hardware implementation, synthesized and did place-and-route, to more accurately estimate hardware complexity.

- DIAZ, I., WILHELMSSON, L., RODRIGUES, J., OLSSON, T., AND ÖWALL, V., "Sign-Bit based architecture for OFDM acquisition for multiple-standards," *NORCHIP*, November, 2009.

  ☞ The author performed Matlab simulations, created a hardware implementation that complied with its algorithm specification and analyzed the hardware costs and benefits.

There is also a list of publications where the author has contributed, yet are not formally part of this dissertation, this list include the following:

- OLSSON, T., CARLSSON, A., WILHELMSSON, L., EKER, J., VON PLATEN, C., AND DIAZ, I., "A reconfigurable OFDM inner receiver implemented in the CAL dataflow language," *IEEE International Symposium on Circuits and Systems*, May, 2010.

- DIAZ, I., SATHYANARAYANAN, B., MALEK, A., FOROUGHI, F., AND RODRIGUES, J. "Highly Scalable Implementation of a Robust MMSE Channel Estimator for OFDM Multi-Standard Environment", *IEEE Workshop on Signal Processing Systems*, October, 2011.

- ZHANG, C., DIAZ, I., ANDERSSON, P., RODRIGUES, J., AND ÖWALL, V., "Reconfigurable cell array for concurrent support of multiple radio standards by flexible mapping," *IEEE International Symposium on Circuits and Systems*, May, 2011.

- ZHANG, C., DIAZ, I., ANDERSSON, P., RODRIGUES, J., AND ÖWALL, V. "Reconfigurable cell array as enabler for supporting concurrent multiple standards in mobile terminals," *10th Swedish System-on-Chip Conference*, May, 2010.

In this thesis is demonstrated that by jointly studying systems, algorithms and hardware architectures, it is possible to develop solutions that mitigate the often contradicting demands of performance, flexibility and energy efficiency. Even though applied to the area of the digital front-end in mobile receivers in wireless communication, the author is confident the approach is valid far outside this field.

This document is divided into several introductory chapters and three parts. The first chapters introduce basic concepts of wireless communication and hardware design needed to understand the scientific contributions. The three parts contain the main contribution areas where the author focused on creating something new and useful for the scientific community.

Part I on page 49 deals with a novel methodology to reduce hardware complexity in low complexity receivers. Part II on page 87 explores the capabilities of highly optimized generic architectures to performed filtering in Long Term Evolution - Advanced (LTE-A), a task so computationally intensive that it is usually done with very specialized hardware. Finally, Part III on page 107 combines algorithm development with a hardware construction techniques in order to deal with the challenges imposed by multi-standard concurrent reception.

Acknowledgments

The number of pages in this thesis is insufficient to name all those who helped me in along this road, nonetheless I will mention a few with hope that my clumsy self does not bedim the important contribution of those not mentioned here.

First, I would like to express my gratitude to my supervisor Professor Viktor Öwall, who always provided his guidance and support as a good leader and true friend, it has been an honor to work for you "jefe", and thanks for singing my paychecks too. My gratitude also goes to my other two great mentors, Professor Ove Edfors and Dr. Leif Wilhelmsson, I learned profoundly much from both of you, I only regret I finish my degree before being able to squeeze-out the last bit of knowledge you both have. I would also like to thank Professor Liesbet Van der Perre and people at IMEC for allowing me to do research in their team as just one more member, especially Lieven Hollevoet, who to this day helps me to stay out of trouble with Belgian authorities.

Thanks to everyone at the EIT department, especially for those in the digital ASIC group, you are all like a family to me and I am happy to have developed ever lasting friendships with each and every one of you, the interesting discussions and joyful moments had nourished me deeply and I am infinitely grateful for that. My appreciation also goes to the secretaries and technical staff who made my life as PhD student smooth and satisfying; it is more pleasant to carry out your work when someone handles the paper-work and technical details.

I would also like to express my appreciation to my Mexican and Norwegian families. My Mexican family always cultivated in me the best values of honesty, hard-work, and perseverance. They were always so assure of some greatness in me to a point where they manage to convince me. And my Norwegian family that accepted me with wide open arms and adopted me as

one of their own, and helped us in ways I couldn't imagine.

Last and most importantly, I want to thank my beautiful wife Maria. We took this project together and this achievement is every bit yours, even more so since you made it possible. To Elias and Helena, "mis hermosos", even though you were not the reason why I started this, you are the reason why I went on finishing it and the reason I go on doing anything else in my life. Without you I would have finished in half the time.

Isael Diaz Palacios
Lund, March 2014.

1

Introduction

Wireless communications is one of the fastest growing market segments. According to estimations by the International Telecommunication Union (ITU), see Figure 1.1, there is today the same number of mobile subscribers as the number of people living on this planet [1]. Furthermore, it is expected to accelerate with the incorporation of the Internet of Things (IoT), where many more devices will be connected wirelessly to the Internet.
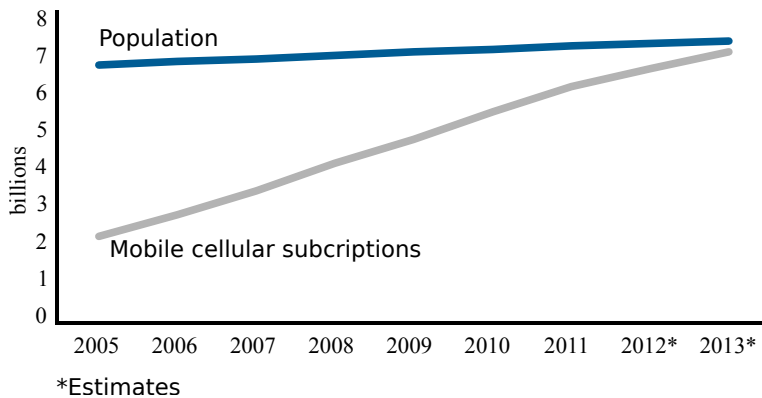
Figure 1.1.: Number of Mobile subscribers in the world [1].

The IoT intends to connect to the Internet each and every device that benefits by any way from being connected. Some estimations predict that we will have 50 billion devices connected to the Internet by the year 2020 [2]. Among the application areas that are expected to benefit from this highly connected infrastructure are, wireless sensor networks, automated highways, machine-

1

to-machine communication, automated factories, smart homes, appliances, remote tele-medicine, etc. There are already some insights on how all these devices will shape our future connectivity. For example, sensor networks will play a major role in connecting trillion of sensors to the Internet. These sensors will be located everywhere and will provide information about everything around us. This unleashed layer of connectivity is commonly called sensory swarm [3]. Even though true wireless ubiquity and sensory swarm are evidently part of our future, there are still major technological challenges that need to be overcome.

Today, these challenges are addressed by two efforts combined. The first consists in making better use of the radio resources by designing better Radio Access Technologies (RATs), and the second consists in designing efficient implementations of radio transceivers in order to take the most advantage of the radio resources with minimum energy consumption.

This thesis deals with the second effort, i.e., efficient design of radio transceivers, and does that by taking a cross-disciplinary approach; i.e., by combining algorithm design and circuit construction techniques to benefit the most from both disciplines.

In order to place into context the contributions presented in this thesis, a number of introductory chapters are included. These chapters deal with important concepts in wireless communication and wireless communication standards. They are not intended to give a detailed description on these significant complicated subjects, but rather, give some reference information on terms and concepts used later in the thesis.

1.1. THE NEED FOR ABSTRACTION

In order to cope with the complexity of designing advanced wireless communication infrastructures, the entire infrastructure is divided into abstraction layers so that each layer is independent and serves as a division between knowledge of specific fields. One of the most common examples of this abstraction is the Open Systems Interconnection (OSI) layer used in network design as a framework for dividing protocols into various layers. This abstraction provides a divide-and-conquer approach to the already complex task of designing a fully operational data network [4]. A brief description of the layers is presented as follows:

7. APPLICATION: This layer is in charge of the interaction with the user. It presents information in a format understood by the user and receives commands to be translated into internal protocols to be passed to lower layers.

6. PRESENTATION: Provides services to the application layer. This is done using a specific semantic the user can interpret independently from the data. Typical tasks within this layer are: data encapsulation, encryption, encoding, *etc.*

5. SESSION: Is in charge of the connection between terminals or computers in the network, determines whether the connection between the communicating parts is two-way or one-way. In the first, both parties transmit at the same time whereas in the latter only one active transmission is allowed. This layer is in charge of opening and closing connection sessions as required.

4. TRANSPORT: This layer is in charge of sending packages from one address to another in the network. It typically adds error control methods to the communication and maintains control over failed and successfully sent packages.

3. NETWORK: This layer defines protocols needed to transfer packages from a point to any other point within the network. Here many packages are clustered within datagrams. Each datagram is given a specific path or route to find its destination within the network.

2. DATA LINK: This layer is in charge of providing a reliable connection between two directly connected nodes. It is also in charge of detecting and correcting errors occurred in the physical layer. It defines protocols to be used in the delivery of information in dependence to the network topology. e.g., point-point, point-to-multi-point, ring connection, *etc.*

1. PHYSICAL: Typically called PHY for short, is in charge of defining the protocol between two directly connected points in the network and the specification of all electrical and physical properties of the connection. The protocol varies in accordance to the transmission medium, e.g., copper, fiber cable, wireless connection, *etc*. This thesis mainly deals with the Physical Layer (PHY) layer, and the trade-offs present with its implementation. While some of the innovations presented impact directly the PHY layer, it might carry out repercussions to higher layers within the OSI model as well.

1.2. MATHEMATICAL MODELING VS. THE REAL WORLD

Further abstraction is found in the way the protocols within each layer are implemented. Specially in the PHY layer where the actual transmission of information takes place via the interaction of physical forces (in radio applications the force in use is electromagnetic). These interactions are modeled as mathematical processes with many ideal (non-realistic) assumptions.

On the one hand, in wireless communications it might be theoretically possible to accurately model all natural processes involved in the wireless transmission. However, this is often not desired, as the actual implementation of such systems is too complex to be done in practice, thus, approximate models are used instead of precise ones.

On the other hand, in circuit design, there are also mathematical models that describe the interaction between the different components that constitute all parts in a circuit, e.g., capacitors, transistors, cables, resistors *etc*. However, as circuits grow larger in size and hence in number of elements, abstractions models are used to handle the large complexity. A common division in circuit design is the partition between analog and digital circuitry. The way each partition works and the way each is designed is completely different.

Due to the large complexity enclosed within each of these abstraction models, there is a well-defined language for each of the disciplines, and it becomes increasingly more difficult for a single individual to be able to master them all. Even though abstractions bring clear benefit in hiding complexity it might also hinder quality, given that possible cross-layer optimizations are set aside or simply ignored. In this thesis the cross-layer optimization is the central object of analysis, more specifically in the realm of mobile terminals.

1.3. HARDWARE IMPLEMENTATION

In the early days of electronic design, circuits were designed for a unique purpose. However, as technology improved, more and more transistors were fit in the same area, to be more precise, in 1965 Gordon Moore predicted that the number of transistors that could be placed in a single die increased exponentially with time [5]. This statement became known as "Moore's law", which was a term coined by Carver Mead, and has driven the electronic industry for the last 50 years.

As the number of transistors in a single chip increased, the semiconductor industry has found ways to connect them in various types of configurations. Each configuration (hardware architecture) represents a different paradigm in terms of energy consumption and flexibility, ranging from the very energy efficient (ASIC) to the very flexible (GPP). Figure 1.2 on the facing page illustrates a general map on how common hardware configurations are compared in the flexibility-energy spectrum. Application-specific instruction-set processors (ASIPs) are architectural heterogeneous devices, hence, they contain various types of sub-architectures, e.g., Very Long Instruction Word (VLIW), Single Instruction Multiple Data (SIMD), and others. Each of the architectures depicted in the figure is briefly described in the following paragraphs.
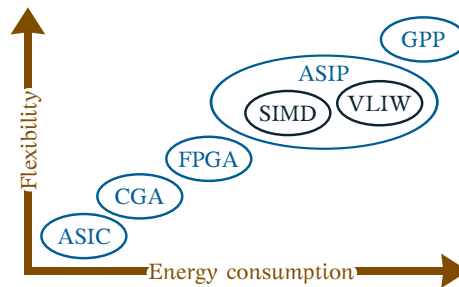
Figure 1.2.: Flexibility vs. energy consumption of popular processor architectures.

ASIC: Application-Specific Integrated Circuits (ASICs) are the customized Integrated Circuits (ICs) for a particular application; they are considered the most energy-efficient alternative for the task they are designed to accomplish. They can be found as stand-alone ICs or as hardware accelerators in System On Chip (SoC) environment. The main drawbacks of ASICs are their lack of flexibility, high fabrication cost, and long development time.

CGA: Coarse-Grain Architectures (CGAs) borrow some characteristics from Field-Programmable Gate Arrays (FPGAs) like, the ability to be programmed and short development time. However, CGAs lack gate-level programmability, which increases their energy efficiency by sacrificing flexibility. A CGA consists of multiple Functional Units (FUs) interconnected via a Network on Chip (NoC), where a FU is a single processing element capable of performing basic arithmetic operations, depending on the intended application more complex operations are sometimes included in the FUs.

FPGA: Field-Programmable Gate Arrays (FPGAs) are digital ICs specifically designed to be configured by the end-user. They consist of a large matrix of bit-level programmable logic blocks, interconnected via flexible interconnections that can also be programmed. Configurable Logic Blocks (CLBs) can be configured to match the desired functionality, that functionality can be reprogrammed while they are placed in their field of action. Among some of the advantages of FPGAs over ASICs are: higher flexibility, scalability, lower price, and shorter development time.

SIMD: SIMDs borrows the concept of simple FU from CGA, but they differ drastically on the way they are programmed and how they are interconnected. As their name implies, all FUs, or processing units, load and execute the same instruction, and the interconnection between FUs is limited. They are

considered vector computers and are extremely beneficial in applications with high degree of data parallelism, like image processing.

VLIW: Very Long Instruction Word (VLIW) architectures are conceptually considered Multiple Instruction Multiple Data (MIMD) computers. They consists of 2 or more parallel processors, and their name is derived from the fact that a single instruction word is long enough to control all processors in the architecture, The burden of task scheduling is shifted to the compiler, which leads to less control logic in the data-path, which can be beneficial in some applications. VLIW architectures offer significant benefits in applications with high level of Instruction Level Parallelism (ILP).

GPP: General Purpose Processors (GPPs) are behind every personal computer, their introduction goes back to the early days of computing. As its name implies, it is designed for a large variety of applications giving GPP very high flexibility but also high energy consumption, i.e., GPPs contain large memories and several type of interfaces, complex arithmetic modules, such as floating point, *etc.,*

1.4. ALGORITHM-ARCHITECTURE CO-DESIGN

The best implementation for any application requires a fine tuning between both algorithm and architecture. Given that design constraints vary considerably among different applications, e.g., silicon area, speed, transmission rate, power, *etc.,* a general abstraction model results in sub-optimal implementations. Thus, efficient implementations can only be realized if both algorithm and architecture are prepared towards a common goal. As an example, consider the fastest wireless transfer record, which is 100 Gbps over 20 m [6]. This achievement was possible thanks to the right combination of theoretical modeling and a precise implementation. Theoretically, the maximum transmission rate depends on the transmission bandwidth, channel conditions, *etc.* However, the theory is only half of the story, even under the most precise mathematical model of the channel, if the hardware realization is imperfect, those perfect conditions cannot be met.

In this thesis, the algorithm-architecture optimization is carried out by the following methodologies:

TAKING THE SIDE OF ALGORITHM DESIGN: Here an application specific algorithm is taken, e.g., signal detection, channel estimation, symbol decoding, modulation, *etc.* The algorithm is designed with basis of transmission performance. Once the algorithm is defined, imperfections originated in the circuit realization are included. Then, proper adjustment to the algorithm

follows in order to account for those imperfections. If the appropriate adjustments were performed, the modified algorithm presents a more adequate alternative than the original. Often the goal is not performance, but energy consumption, die size or other. In that case the adjustments have to be done according with that specific goal in mind so that the final result is better than the original in terms of a specific design parameter; performance is traded for energy consumption.

TAKING THE ARCHITECTURE SIDE: In modern mobile terminals it is challenging to arrive at an optimal implementation. This is due to the fact that most modern terminals are multi-purpose. Under these circumstances it is attractive to look into the hardware side for optimization opportunities. Modern architecture configurations such as VLIW, Coarse Grain Reconfigurable Architecture (CGRA), SIMD, provide a high degree of flexibility as they exploit parallelism inherent in the algorithm or application. Then by selecting an appropriate combination of these configurations, it is possible to increase: usability, cost-effectiveness, platform life-cycle, *etc*. Therefore, in this methodology a group of applications or a single multi-scenario application is chosen, then a combination of architectures is selected that optimizes a specific design criteria, e.g., area, speed, transmission rate, power, cost-effectiveness, platform life-cycle, *etc*. As a result, the final architecture represents a better alternative in terms of that optimization constraint, than multiple highly specialized circuits designed for all possible cases.

2

Wireless communication basics

The work done in this thesis applies to one type of wireless communication, namely wide-band digital communication with Orthogonal Frequency Division Multiplexing (OFDM) technology. More specifically, it applies to initial signal acquisition and filtering in the receiver terminal, which are parts of what is often referred as Digital Front-End (DFE). However, the underlying concepts behind algorithm-architecture co-design can find application in many other disciplines. In in this chapter a general introduction of wireless communication is given, in order to place the relevance of this thesis in a larger context.

The basic principle of wireless communication consists of transmitting information from an information source, referred as Transmitter (Tx), to an information destination, referred Receiver (Rx), via a wireless propagation channel. The Tx can modify the information to make it better suited for reception. Then it can make better use of the transmission medium by using modulation and can also protect it from unwanted listeners by encryption. It can make it more resistant to errors by adding coding. Finally, the Tx can format it so that multiple Rxs can extract only portions of it. Conversely, the Rx needs to perform the inverse transformations in order to interpret the received data; it also needs to correct for any error introduced by the propagation channel.

In wireless communications one of the main challenges remains in the propagation channel, and there are three fundamental aspects that make communication difficult. The first is the fact that the propagation channel properties vary with time, known as fading. The second is the fact that there is no physical connection between both transmitting ends, resulting in significant interference and synchronization issues during the communication [7]. The third is the unavoidable imperfections in practical implementations resulting in additional noise.
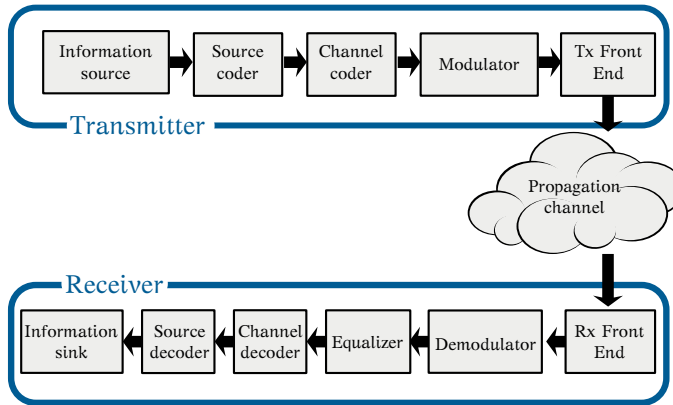
Figure 2.1.: Block diagram of a typical transmission link.

2.1. FUNCTIONAL MODEL OF WIRELESS COMMUNICATION

The communication link can be abstracted in a functional model which involves generation of information, its preparation for transmission, the actual transmission over the channel, reception, correction of received information, and its final interpretation at the receiver end. These abstractions are illustrated in Figure 2.1, where a functional block diagram of a typical communication link is illustrated in several blocks.

There are more modules in a real transmission system that are not illustrated in this diagram. Among these modules are, those for packaging, pilot insertion, *etc*.

In the diagram, each block represents a data transformation. Various data transformations are necessary to increase the chances of the information being successfully interpreted by the Rx. Some transformations are directly related to the properties of the information to be transmitted and other related to the characterization of the transmission medium, or propagation channel, as it will be clear in the remainder of this section.

2.1.1. INFORMATION SOURCE/SINK

The information source and sink represent the origin and destination of the intended data. In principle the information to be transmitted can be anything, and typical examples include, voice, music, image, video, or any other form of digital media.

2.1.2. SOURCE CODER/DECODER

The source coder typically uses known properties of the signal in order to remove redundancy and thus, reduce it to a minimal representation. among such coders there are speech coders, image/video coders, audio coders, *etc*.

2.1.3. CHANNEL CODER/DECODER

The channel coder introduces a structure redundancy that can be used on the receiver side to minimize errors. Many types of codes are available and it is still an active research field. Some examples are, block codes, convolutional codes, turbo codes, and intricate types such as Low-Density Parity-Check Code (LDPC).

2.1.4. MODULATOR/DEMODULATOR

The modulation consists of mapping information bits to signal waveforms that carry the desired information in a format better suited for transmission. Conversely, the demodulation does the opposite operation. However, since the transmission medium influences the received waveform, an additional equalization step is required at the receiver.

Each waveform can carry from a single to any number of bits, with the main limitation that as the number of bits increases, the probability of losing some information is increased. The main motivation in choosing a specific modulation is to transmit as much information as possible over a channel with a specific bandwidth, while at the same time satisfying a given probability of error, better known as Bit Error Rate (BER).

A list of common digital modulations techniques include among others: Pulse Amplitude Modulation (PAM), Phase Shift Keying (PSK), Frequency Shift Keying (FSK), and Quadrature Amplitude Modulation (QAM). Special attention is given to QAM, since it has been widely adopted in cellular and Local Area Networks (LAN) systems, as this type of modulation is capable of enclosing a large number of bits per Hz. In QAM the information is placed as amplitude values on two orthogonal signals (in-phase and quadrature), typical realizations include 4QAM, 16QAM, 64QAM, and 256QAM. Figure 2.2 on the following page depicts a constellation diagram for 16QAM with Gray coding[1].

2.1.5. EQUALIZER

When waveforms carrying the desired information are exposed to the transmission medium, e.g., a wireless medium, they are unavoidably affected by it, making difficult the correct interpretation of the received information. An equalizer can be used to remove the effects introduced by the channel; this

---

[1]Gray coding is a binary representation designed so that neighboring constellation point differ only by one bit.
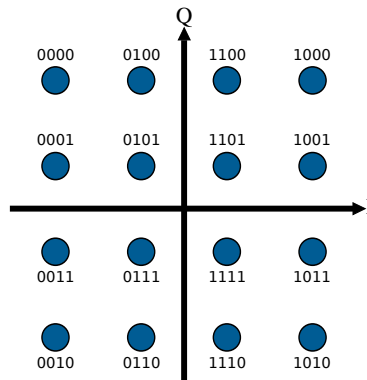
Figure 2.2.: Constellation diagram of rectangular 16QAM with Gray coding.

is done by relying on various equalization techniques. In some cases, as typically done in OFDM, the equalization stage is divided in detection (estimation) and correction (equalization).

In general, equalizers can be characterized by type, structure and algorithms [8]. There are countless equalization methods, and it is to this day, a very active field of research.

2.1.6. TX/RX FRONT-ENDS

The Tx front-end has the function of taking a baseband signal and placing it at a carrier frequency. Then, it follows that the Rx does exactly the opposite operation. Modern front-ends are typically divided in an Analog Front-End (AFE) and a DFE, where the limits of what operations belong to what domain (analog or digital) are not well defined, and it's closely related to specific transceiver architecture. In this thesis, a special focus is given to the DFE, especially when operating in connection to Software Defined Radio (SDR) architectures.

2.1.7. PROPAGATION CHANNEL

The propagation channel in combination with the receiver's noise level limits the ultimate transmission efficiency. Wireless communication channels are characterized by multi-path propagation. As the signal travels from Tx it is reflected, diffracted and scattered by structures encountered on its path towards the Rx forming what is denominated as multi-path components. These components are added at the receiver end, and need a special treatment.

An illustration of the multiple path transmission concept is shown in Figure 2.3 on the next page. In the figure, there a signal bouncing off two
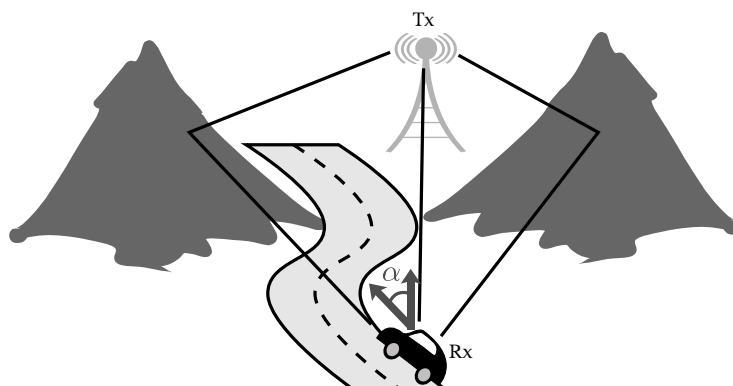
Figure 2.3.: Muti-path components present in wireless communication.

mountains, and hence creating two multi-path components together with a direct line of sight component. Another typical characteristic of wireless channels is the presence of Doppler effect, due to the aforementioned multi-path components, and the mobility of either Tx or Rx. The relative motion produces frequency shifts on the transmitted signal. This is illustrated in Figure 2.3. Doppler effect can be an important source of imperfections in the transmission link. More about the propagation channel and Doppler effects in Section 2.3 on page 20.

2.1.8. ANTENNA TECHNIQUES

Multiple-antenna techniques can be used to increase transmission data-rate or to improve reliability. More specifically, it is possible to utilize multiple antennas for: a) beam-forming, b) increase diversity, c) suppress interference, and d) spatial multiplexing. The improvement takes place without any increase of transmission bandwidth. However, these techniques have a significant cost in terms of hardware and computational requirements in both Tx and Rx. Possibly the simplest technique, from the implementation perspective, is the use of the additional antennas to provide antenna diversity. Since different versions of the same signal are received by several antennas, the receiver can choose which antenna to listen to based on its associated channel quality.

Multiple antennas systems can be classified in four categories: Single Input Single Output (SISO), Multiple Input Single Output (MISO), Multiple Input Multiple Output (MIMO), and Single Input Multiple Output (SIMO), depicted in Figure 2.4 on the following page.

a) SISO   b) MISO   c) MIMO   d) SIMO
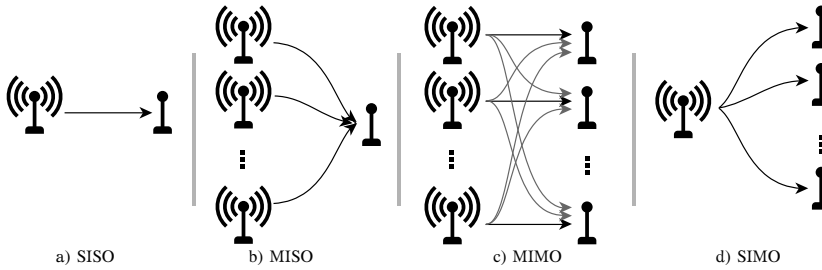
Figure 2.4.: Four different types of antenna configurations, a) SISO, b) MISO, c) MIMO, and d) SIMO
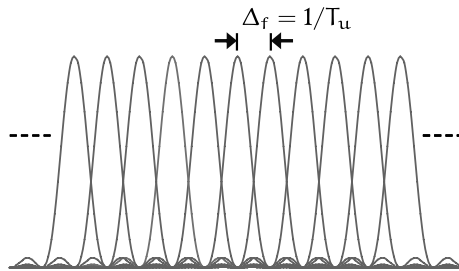
Figure 2.5.: Frequency domain bandwidth occupied by an OFDM system with a sub-carrier spacing Δ_f, specifically designed to maintain orthogonality.

2.2. ORTHOGONAL FREQUENCY DIVISION MULTIPLEXING

OFDM is a modulation technique that consists of placing information on multiple carrier frequencies (sub-carriers, or tones). This is particularly beneficial for high data rate transmission in delay dispersive environments. OFDM converts a high data rate transmission into several parallel lower data rate transmissions. The lower data rate transmissions are orthogonal and can be easily equalized.

The sub-carriers are placed at the minimum allowed frequency separation needed to maintain orthogonality. The spectral efficiency of such implementation is rather high compared to other communication technologies. This can be shown in Figure 2.5, where the sub-carrier spacing Δ_f equals the inverse of the symbol timing T_u; necessary to ensure the orthogonality.

Various modern RATs have adopted OFDM technology as the modulation technique at the core of the standard specification. The popularity of OFDM can be traced to one or more of the following benefits:

- high spectrum efficiency, resulting from dense sub-carrier packing,
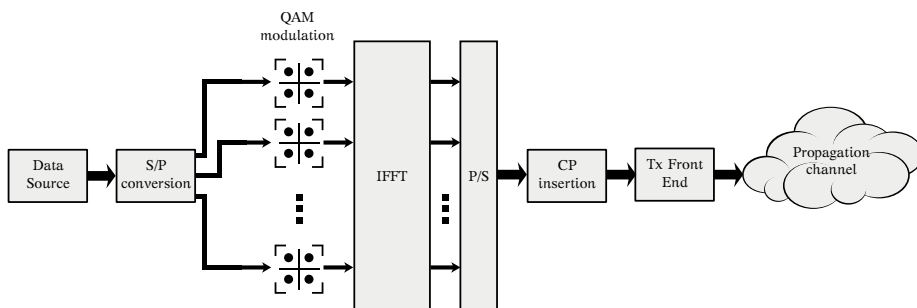
Figure 2.6.: Block diagram of an OFDM transmitter.

- it can operate on highly frequency selective channels (delay dispersive environment), and

- the simplification of equalization techniques in the receiver terminal which is typically provided with limited resources.

In the remainder of this chapter some basic concepts regarding OFDM technology are discussed. However, it is important to note that only the aspects affecting the work in this thesis are addressed, and others are intentionally left out. If reader feels the need of acquiring deeper knowledge regarding the subject, they are encouraged to read [9] [10], [11], [12], and [13].

2.2.1. OFDM TRANSCEIVER

As previously discussed, the key idea behind OFDM is to divide the spectrum into several narrow-band channels, in order to make each narrow-band channel immune to frequency selectivity, i.e., each sub-carrier experiences a quasi-flat fading channel. This is typically done with the aid of Inverse Fast Fourier Transform (IFFT) and Fast Fourier Transform (FFT) modules sitting in the transmitter and the receiver, respectively. A block diagram illustrating the most important parts of an OFDM transmitter is depicted in Figure 2.6, where Serial to Parallel (S/P) and Parallel to Serial (P/S) are purely conceptual as they don't affect the actual information in any manner.

The QAM modulation phase in the figure, corresponds to the modulation introduced in Subsection 2.1.4 on page 11. In OFDM the modulation used is typically restricted to QAM and PSK [11]. A Guard Interval (GI) filled with a Cyclic Prefix (CP) is used in order to avoid Inter-Symbol-Interference (ISI) and Inter-Carrier-Interference (ICI). However, as it will be seen in Subsection 2.2.2 on the following page, it leads to a loss of transmission efficiency.

As one can expect, the receiver applies the inverse operations of all the depicted modules in reverse order, i.e., Front-End (FE) → CP removal → S/P

$\rightarrow$ FFT $\rightarrow$ QAM demodulation $\rightarrow$ P/S $\rightarrow$ data sink. It is important to note that the OFDM transceiver chain is easily placed in a communication link system perspective, this can be visually conceived by replacing the Modulator/Demodulator block from Figure 2.1 on page 10, with the OFDM transceiver chain without duplicating the FEs.

The utilization of FFT modules is an efficient way for placing the modulated symbols onto each sub-carrier. There are two possibilities to map the sub-carriers onto the number of available FFT points: First, if the number of sub-carriers is not a power of two, one can leave zero sub-carriers at the boundaries of the FFT and place information only on the middle FFT points. This is done in several RAT specifications, e.g., Long Term Evolution (LTE), IEEE 802.11 (WiFi), among others, since it leads to simpler reconstruction/transmission filters. Second, is to use a mixed-radix FFT with equal number of points as the number of sub-carriers available. This solution is less efficient from the perspective of hardware implementation, but it might be beneficial in some well defined scenarios [14] and [15].

2.2.2. CYCLIC PREFIX

The Guard Interval (GI) extends the duration of the OFDM symbol to allow the channel's impulse response to settle during the duration its duration in order to avoid ISI. This is also graphically depicted in Figure 2.7, where it can be seen that the last part of the symbol is copied to the start of the GI, just before the OFDM symbol. This is done in the time domain and it is referred as CP insertion.
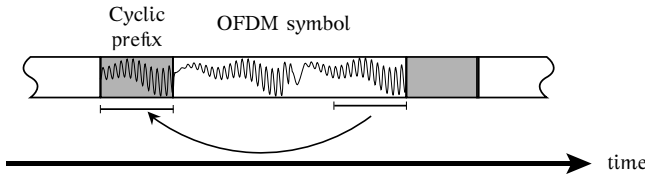
Figure 2.7.: Typical OFDM symbol with its corresponding CP.

ISI is the undesired interaction between consecutive OFDM symbols due to the influence of the propagation channel to the transmitted signal. The propagation channel can be modeled as a Linear Time Invariant (LTI) system[2] with a limited memory, and therefore, a convolution with the channel produces the transient period at the beginning and end to each OFDM symbol. The duration of this period is given by the length of the channel's impulse

---

[2]In reality, the wireless propagation channel is time invariant only for some special cases. However, the channel can be considered quasi-stationary for a short period of time, given by the relative motion between Tx and Rx, read more about channel modeling and its properties in Section 2.3 on page 20.
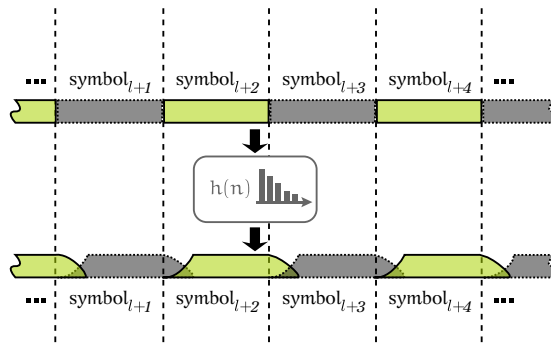
Figure 2.8.: OFDM symbols that lack Guard Interval. The dark intersection area corresponds to that ISI introduced by the propagation channel h.

response. This is illustrated in Figure 2.8, where it can be seen the ISI as the intersection between the transient of the previous symbol with the transient of the current one.

Then, by removing the part of the CP at the receiver, given that the CP is larger than the channel's impulse response, ISI can be completely mitigated. However, a careful choice for the duration of such CP has to be made, since an increase of the duration of CP results in loss of spectral efficiency. The CP has another important contribution, since it allows the equalization to become a simple multiplication per sub-carrier in the frequency domain.

FREQUENCY ERRORS IN OFDM

There are two types of frequency misalignments in OFDM, both resulting in undesired ICI. First, a frequency offset between local oscillators causing a slight shift in the carrier frequency, more commonly referred as Carrier Frequency Offset (CFO). Second, a frequency offset between master clocks, at both Tx and Rx which results in frequency misalignment between sub-carriers, this misalignment is commonly known as Sampling Frequency Offset (SFO).

ICI is the loss of orthogonality between sub-carriers. In order for OFDM to function properly, the transmitter and receiver sub-carriers must be perfectly aligned, and the FFT modulation has to be applied to exactly a complete symbol period. The presence of CFO breaks the sub-carrier alignment from transmitter and receiver, even though the sub-carrier spacing is the same at both ends. Furthermore, a SFO contributes by an apparent scaling of the spectrum as seen by the receiver.

ISI is illustrated in Figure 2.9 on the next page, where it is seen that the OFDM system with ICI contains a frequency separation of Δ_f. In the sub-figure to the right the desired sub-carrier spacing plus an error ϵ due
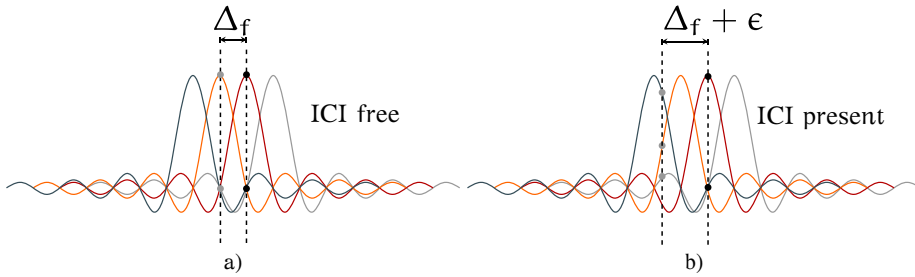
Figure 2.9.: OFDM frequency bandwidth a) without ICI, and b) with ICI due to frequency misalignments.

to misalignment. As a consequence, the point is not sampled at the zero crossing, opposite to what is seen in the sub-figure to the left, where there is no ICI, i.e., the sub-carrier spacing is exactly Δ_f, at the sampling point. ISI due to frequency misalignments drastically affect the perceived Signal to Interference Ratio (SIR) and need to be eliminated. In this thesis, the CFO estimation and removal is addressed and studied in some detail.

2.2.3. SYNCHRONIZATION

Synchronization is a crucial process in any digital communication scheme. The need for synchronization arises from the fact that Tx and Rx do not possess a common time nor frequency reference.

When a mobile terminal is initiated, it needs to search for a signal that announces the presence of a base station in the proximity. This process is typically called *acquisition* [16], and it is one of the important procedures in the establishment of a successful communication link. Typically during *acquisition*, the mobile terminal performs synchronization, i.e., symbol-timing detection, initial CFO estimation, CFO correction, and initial channel estimation. In this thesis the synchronization process is given special attention, more particularly the part of acquisition. The methods and architectures presented later in this document are designed to estimate, among other parameters, symbol timing, and CFO, during acquisition stage.

Synchronization in OFDM can tolerate relatively large errors in the symbol-timing estimation in comparison to a single-carrier system [11]. However, in contrast to single-carrier systems, it is more sensitive to frequency misalignments; due to the overlapping of narrow-band spectrum in OFDM.

Two types of synchronization methods can be found in literature, namely, data-aided and non data-aided (blind). The first type uses information embedded in the signal for this particular use, whereas the second relies on the CP and typically suffer from lower performance, since the CP is affected by ISI.

There are many methods in literature for aided and non-aided synchronization. Some examples of aided synchronization include Maximum Likelihood (ML) estimation [17], and some customizations of this given in [18–20]. Among the non-aided methods, one of the most commonly cited is the one using the CP [21]. In this thesis, the author explores both types, but gives special attention to non-aided, as in some cases can also be used to estimate other parameters, such as Signal-to-Noise Ratio (SNR).

2.2.4. SPECTRUM IN DUPLEX ARRANGEMENT

A duplexing arrangement is not confined to OFDM systems, neither to wireless communication, and it has wide application among all types of communication. Full duplexing allows communication to happen in two directions simultaneously, i.e., from Tx to Rx, and vice versa. In wireless communication the most logical partition when duplexing is in time and/or frequency. Thus, two obvious multiplexing configurations are Frequency Division Duplexing (FDD) and Time-Division Duplexing (TDD), and some hybrid in between, like half duplex FDD. These are depicted in Figure 2.10, and briefly described as follows.
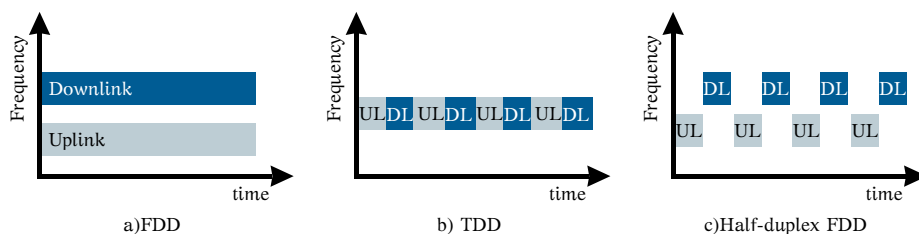
Figure 2.10.: Frequency and time division multiplexing.

FREQUENCY DIVISION DUPLEXING, is the duplexing scheme where the entire transmission bandwidth is divided into two completely separate transmission bandwidths, one for Uplink (UL), and one for Downlink (DL). The advantage of such a configuration is that communication between both ends can happen simultaneously. However, it requires a paired spectrum and it imposes complexity in the receiver, as it needs completely independent and well isolated Tx/Rx chains with no straight-forward hardware reuse.

TIME DIVISION DUPLEXING, is the scheme where the transmission of UL or DL is done in turns. Contrary to what happens in FDD, this scheme does not require paired spectrum and thus, it is easier to allocate in a fragmented spectrum. As a consequence, this is scheme is the preferred choice for many operators.

HALF DUPLEX FDD, is a duplexing scheme where there is an exclusive part of the spectrum of each transmission and reception, and these two operations do not take place at the same time, this is typically used to reduce complexity at the terminal,i.e., many hardware blocks can be reused, and it is supported by several RATs, e.g., Global System for Mobile communication (GSM) and LTE.

2.3. THE WIRELESS CHANNEL

Many things can be said about wireless channels, they can be classified by many criteria e.g., based on representation, modeling-technique, communication topology, statistical parameters, *etc.* The author has no intentions into discussing details about wireless communications channels, but rather, just to put into context the application field of the work presented. For that reason some important concepts are briefly described in this section.
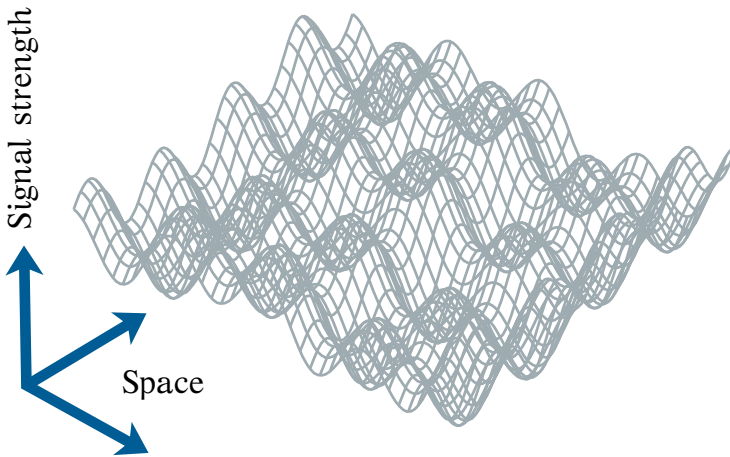
Figure 2.11.: A two-dimensional sine wave propagating in a three multi-path components environments, the three reflected signals super imposed and form a complicated wave in space.

The wireless channel is one of the most important elements in the wireless communication scheme. The channel in combination with noise set the ultimate limit of the communication performance. Contrary to a wired channel, in wireless channels there are multiple propagation paths that can contribute constructively and destructively to the received signal, i.e., the transmitted signal is reflected, diffracted, or scattered along each multi-path on its way to the receiver. This is illustrated in Figure 2.11 where the superposition of three reflected paths are depicted, note that the resulting superposition does

differ from a uniform two-dimensional sine wave propagation, as the phases of the reflected copies and its amplitudes disturb each other.

There are, however, other scenarios where there is no multi-path components to disturb the transmitted signal. Those scenarios are governed by *free space attenuation*, where the received signal strength is a function of the communication distance.

There are channel modeling techniques that take into consideration many elements in the propagation path of the signal, e.g., reflection, scattering and diffraction and resolve many multi-path components. However, these methods are often very complex as they need specific knowledge of the electric properties of materials involved in the transmission. Furthermore, considering that in a real scenario many objects move, e.g., cars, people, even the hand that holds the receiver terminal. Therefore, it is often chosen to describe the channel as an stochastic process with certainties and probabilities.

2.3.1. STATISTICAL DESCRIPTION OF THE CHANNEL

Due to the large complexity required to resolve all individual multi-path components, it is often desired to describe a channel statistically. This implies that the channel is fully characterized not by individual channel realizations but by its full statistics. As a consequence the received power variation are classified in two categories, path-loss and fading. Briefly described as follows.

PATH-LOSS

The Path-loss is an unavoidable loss of signal strength due to the distance the transmitted wave has to travel to reach the receiver. It also depends on the transmission frequency, and is closely related to *free space attenuation*, mentioned on this page, and be seen in Figure 2.12 on the next page as a straight line with negative slope.

FADING

Fading describes the variations of the signal as seen from the receiver. Fading varies in relation to time, transmission frequency, and geographical position. Strong interference is typically known as a *deep fade* and it impacts negatively the communication link by a severe drop of the received SNR.

Fading is typically modeled as a stochastic process that varies in amplitude and phase of the transmitted signal. It is commonly characterized by large-scale and small-scale fading, visually depicted in Figure 2.12 on the following page.

- large-scale or macroscopic fading, is the result of the environment changing due to motion. Its behavior is typically attributed to shadowing of large objects, e.g., buildings, mountains, *etc*. Experimental
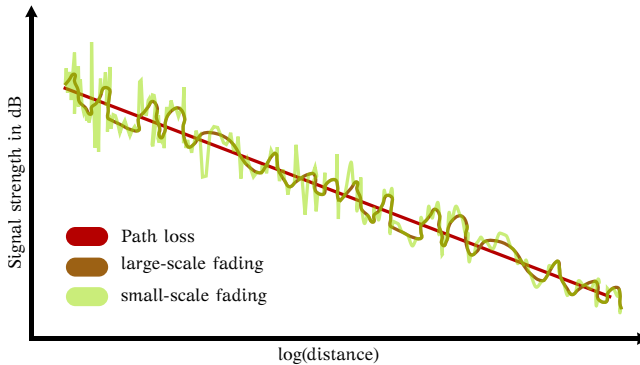
Figure 2.12.: Fading and its three parts, path-loss, large-scale, and
small-scale fading.

results show that fading can be modeled by a log-normal distribution;
i.e., the amplitude in dB follows a Gaussian distribution.

- small-scale or microscopic fading, refers to signal strength variation
over very short distance (approximately around one wavelength), and
these fluctuation are mainly attributed to multi-path components. Small-
scale fading is mathematically modeled as a Rayleigh distribution if
there is no dominant component present, commonly known as Line of
Sigth (LOS), or as a Rician distribution in case of an existing LOS
component.

DOPPLER SPECTRA

When either Tx, Rx, or something in the environment are moving during
transmission, the various directions of the multi-path component produce
different frequency shifts on the received signal. As a consequence, the
spectrum perceived by the receiver is shifted. The typical visualization of
this phenomenon is the transmission of a single sine wave over a narrow
channel, where there is infinite number independent and identically dis-
tributed (i.i.d.) isotropic multi-path components. Under this assumption, the
resulting Doppler spectrum has a bathtub shape, and it is delimited by the
maximum Doppler shift (v_D), illustrated in Figure 2.13 on the next page.

Note that the Power Spectral Density (PSD) is infinite at the maximum
Doppler shift. However, remember that this is an illustration of an ideal as-
sumption never true in practice. The maximum Doppler shift is an important
parameter since it describes the *frequency dispersion*. It can also serve as an
indicator of *temporal variation*, indicating to what extent a channel can be
considered time-invariant.
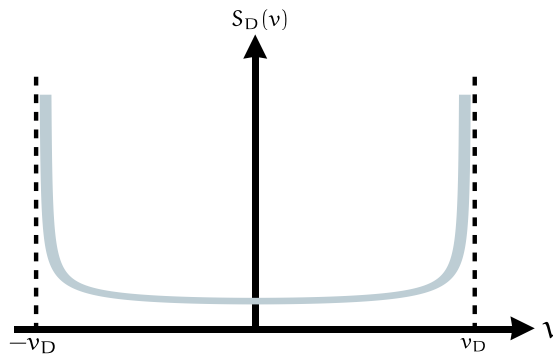
Figure 2.13.: Illustration of ideal Doppler spectrum with infinite number of multi-path components arriving from all possible directions to the receiver.

CHANNEL ASSESSMENT

A channel can be generalized based on two pairs of parameters, namely, *delay spread-coherence bandwidth* and *coherence time-Doppler bandwidth*. They provide a very general idea on how a channel behaves during most transmissions, thus they can be used as tools to intuitively assess a particular channel. Note that, this is a very generalized simplification and only serves as an intuitive way to asses wireless channels.

- *Delay spread*, describes the approximate duration of the impulse response of the channel, and it is connected to the *frequency selectivity*. Frequency selectivity mainly depends on the environment and the transmission bandwidth, since narrow-band systems suffer more from selectivity than wide-band systems [9].

- *Coherence bandwidth*, is often used as an indication on how stable the channel is in the frequency domain, the larger the coherence bandwidth the flatter the channel. It maintains an inversely proportional relation to the *delay spread*.

- *Doppler bandwidth*, is the bandwidth given by the *Doppler spectra*, already described on the facing page. It is connected to the environment and the motion, larger speeds result in larger Doppler bandwidth. It is inversely proportional to the channel's *coherence time*.

- *Coherence time* is a parameter that tells for how long the channel can be considered static, i.e., time-invariant. A large delay spread results in a small coherence bandwidth, and consequently the channel is perceived as "frequency selective", or "frequency flat" if the opposite applies.
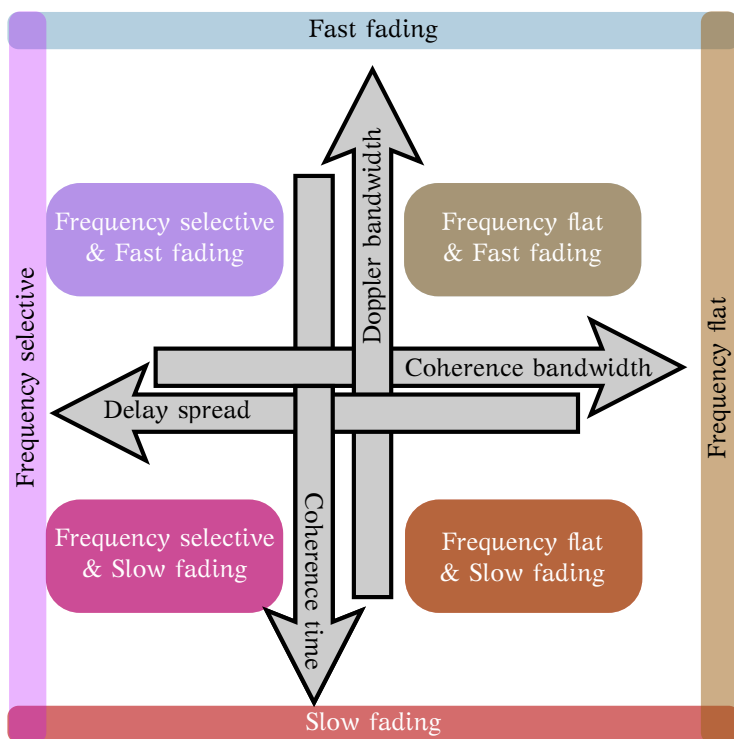
Figure 2.14.: Channel assessment in relation to, coherence time, Doppler bandwidth, coherence bandwidth, and delay spread.

It can be said that if a channel has a large Doppler bandwidth, it also contains a short coherence time, and as a consequence can be categorized as "fast-fading". If the opposite is true, the channel is of "slow-fading" nature. These relations are illustrated in Figure 2.14.

3

Common wireless technologies

Wireless communication is so essential in today's typical life that it can be found virtually anywhere, in stationary and mobile applications, e.g., two-way radios, mobile phones, computers, personal assistants, garage openers, computer accessories, headphones, television, cordless phones, *etc*. The list just keeps growing. In order to enhance device compatibility, interoperability, safety and quality, standardization bodies have defined a set of wireless standards. These standards specify the implementation rules of devices and systems for a particular purpose.

The work performed in this thesis applies to OFDM standards only. More specifically to IEEE 802.11 (WiFi), Long Term Evolution (LTE), and Digital Video Broadcasting - Handheld (DVB-H). This chapter tries to place them into context and highlight their noticeable differences. There are many details intentionally left out as they are not considered relevant to the work presented in later chapters.

3.1. APPLICATION-BASED CLASSIFICATION

Since the number of applications is large and some requirements often contradict with others, a set of standard families is created with similar goals in mind. When it comes to wireless communication four large commercial categories can be identified, namely, Wireless Wide Area Networks (WWANs), Wireless Local Area Networks (WLANs), Wireless Personal Area Networks (WPANs), and satellite communication, briefly described.

3.1.1. WIRELESS WIDE AREA NETWORK

The standards belonging to WWAN category are designed to give service to large physical areas. Broadcasting services like radio or television together
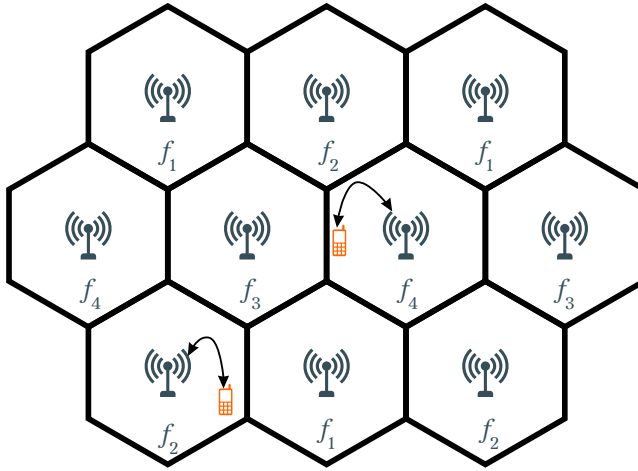
Figure 3.1.: Layout of a cellular system network with frequency
reuse factor of 1/4.

with cellular systems are some examples. Whereas radio and television pro-
vide service in a single direction, cellular systems can provide two-way com-
munication with regional, national, and international coverage. Note that
this is possible by the combination of wireless and wired communication
combined.

Cellular networks are named as such given their capacity to layout net-
works based on coverage cells. The concept was invented by Douglas H. Ring
at Bell Labs in 1947, to be later patented in the 1972 [22]. Cellular networks
takes advantage of the fact that signal power is attenuated by distance, it
is possible to reuse frequencies in separate locations. Figure 3.1 illustrates a
cellular network layout.

Cellular systems are usually refereed as interference limited systems. When
a phone communicates, it does so via a Base Station (BS) servicing the cell
where the phone is located. The energy from other cells with the same
frequency configuration are perceived as interference, more formally named
inter-cell interference. Modern cellular standards may have a heterogeneous
deployment, with micro-cells and nano-cells with shorter coverage embedded
within a BS with larger coverage; typically called macro-cell. In heteroge-
neous deployments the inter-cell interference becomes more relevant. Typical
examples of cellular systems include all standard defined by 3rd Genera-
tion Partnership Project (3GPP), i.e., Global System for Mobile communi-
cation (GSM), Wide-band Code Division Multiple Access (WCDMA), High
Speed Packet Access (HSPA), Long Term Evolution (LTE), and its newest
release Long Term Evolution - Advanced (LTE-A). There are a few other
standards that had evolved from WLAN into offering WWAN services, like

Worldwide Interoperability for Microwave Access (WiMAX) and city-wide WiFi.

3.1.2. WIRELESS LOCAL AREA NETWORK

The goal of WLAN is to act as the wireless counterpart of the more traditional wired LAN topology. The area to be covered is typically a campus, a building, an office, or a house. Users connecting to a WLAN are expected to be semi-stationary, typically sitting at a desk, or moving at pedestrian speeds.

An important factor for the success and rapid adoption of WLAN standards in almost all modern commodities is the fact that these type of standards use the unlicensed spectrum, which saves costs and avoids licenses. However, for exactly the same reasons, there is a good deal of other technologies using the same spectrum causing a great amount of interference.

Institute of Electrical and Electronics Engineers (IEEE) regulates the standardization 802.11, also known as WiFi. There are substantial differences between each member of the family. e.g., the first release of 802.11 was based on Direct-Sequence Spread Spectrum (DSSS), Frequency-Hopping Spread Spectrum (FHSS), and infrared access technologies, whereas newer versions use OFDM.

The evolution of the WiFi standards has brought higher data-rates, better coverage, and an increased level of security. As an example, the latest release of the family is IEEE 802.11af[1] which provides means to access the white space spectrum (this modern technique is known as cognitive radio), and supports data rates up to 568.9 Mbps [23]. Even higher data-rates are expected with future releases.

3.1.3. WIRELESS PERSONAL AREA NETWORK

The purpose of WPAN devices is to connect devices focusing on a single individual and her surroundings. The application scope is localized within a limited space and due to the typical short-range coverage and low data-rate requirements, these technologies can operate on a tight energy budget.

Besides the low-energy consumption, another typical advantage is the low price, due to the operation of such devices on unlicensed spectrum bands. Some examples of these standardized technologies include, Infrared Data Association (IrDA), Bluetooth, Zig-Bee, some versions of WiFi, Near-Field Communication (NFC), and many others.

There also exists a number of non-standardized technologies operating on the same unlicensed bands that can be categorized as WPAN technology, e.g., cordless phones, radio control modules, *etc.* These are typically developed in-house for a very specific purpose and do not follow any particular standardization body, and thus, not compatible with other vendors.

---

[1]Scheduled to be released in March 2014
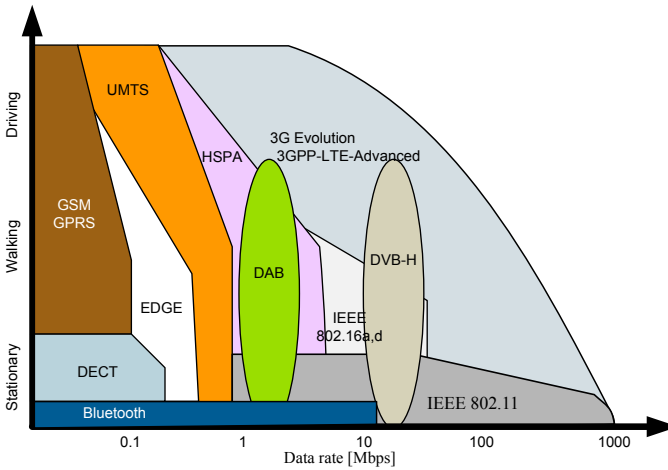
Figure 3.2.: Data Rate transmission vs. Mobility for wireless standards extended from IMEC Scientific Report 2007 [24].

3.1.4. SATELLITE COMMUNICATION

Satellite communication describes the type of communication taken place between an artificial satellite orbiting the earth and one or many devices on the surface of the earth. The intended applications for satellite networks are either broadcast services, such as television, radio, Global Positioning System (GPS), *etc.,* or point-to-point communication. Broadcast services are the ideal application since a satellite can cover a large area and the transmission cost can be shared between large numbers of users. Point-to-point communication, on the other hand is expensive and it is typically confined to inaccessible areas where no other type of communication is possible.

3.1.5. DATA RATE VS. MOBILITY

RATs can also be classified in relation to their data-rate and mobility. Figure 3.2 depicts some of the most popular standards classified in this manner. On the bottom-left corner are the standards that provide the lower data-rate under stationary conditions, which usually results in low energy implementations. Different standards cover different areas in the plot that also place different requirements on the terminals supporting them.

Note from the figure that according to this classification the standards that occupy a larger portion of the plot are more flexible, as they can both provide various data-rates and multiple mobility capabilities. There is, of course, an implicit cost in energy consumption not depicted here, e.g., WiFi operating on low data rate consumes more energy that standard Bluetooth .
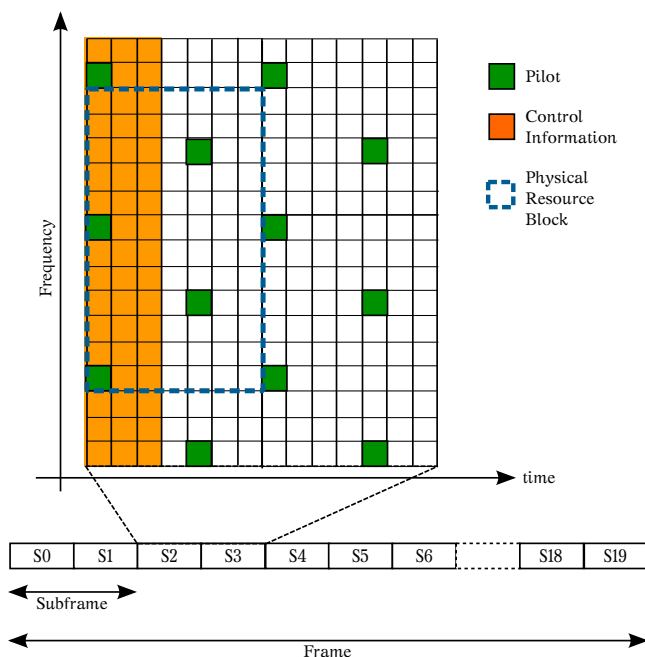
Figure 3.3.: LTE-FDD, DL time-frequency grid with one RB high-lighted

3.2. 3GPP LONG TERM EVOLUTION

This section provides a brief introduction to the physical layer procedures and characteristics of LTE during Downlink (DL), more detailed information about this particular RAT can be found in references [12] [25] [26].

3.2.1. PHYSICAL LAYER PARAMETERS.

The Physical Layer (PHY) is one of the abstraction layers defined by LTE's radio protocol architecture, PHY is in charge of coding/decoding, multi-antenna mapping, channel estimation, synchronization and other functions. The PHY provides services to the Medium Access Control (MAC) layer via transport channels.

In LTE a physical channel is modeled as a two-dimensional grid where the dimensions correspond to time and frequency, graphically represented in Figure 3.3. Each small square represents a single sub-carrier at a specific OFDM symbol and it is referred as Physical Resource Element (PRE). A Physical Resource Block (PRB) is the minimum amount of time-frequency resources that can be assigned to any particular user and consists of 12 sub-carriers within 7 symbols, this accounts for a total of 84 PREs.

In the time-frequency grid there are reference symbols (pilots), used for channel estimation purposes, the figure on the previous page shows only one layer for antenna port 0. LTE features MIMO antenna techniques supporting up to 8 antennas. Thus, in reality, there is one such grid per antenna port. The position of the pilots change depending on the antenna ports. The pilots on one antenna most remain silent in all the others in order to avoid interference at pilot positions.

Note also that the first three OFDM symbols in each sub-frame, excluding pilot positions, are reserved for control information. An LTE BS uses these resources to broadcast important information to all users, e.g., number of assigned Resource Blocks (RBs), modulation used, coding used, *etc*. It is important to mention that this control information spans the entire transmission bandwidth regardless the load of the network. However, the number of OFDM symbols used for control information varies from 1 to 3.

In the time domain, the physical layer is divided into frames, where each frame is 10 ms long and contains 10 sub-frames of 1 ms each. A further partition is done in slots; each sub-frame contains 2 slots of 0.5ms each. The final partition is into OFDM symbols; each slot contains 7, 6, or 3 symbols, in normal CP, extended CP and Multimedia Broadcast Single Frequency Network (MBSFN) mode, respectively.

LTE is an evolution from previous 3GPP standards. Thus, it is reasonable to expect that, for compatibility and historic reasons some parameters are the same or related to those in other members of the 3GPP family. For example, the sub-carrier spacing Δ_f is selected as 15 KHz, which is $1/2^8$ of 3.84 MHz, the chip-rate defined for Universal Mobile Telecommunications System (UMTS). Thus in principle, a terminal can have the same clock reference to interact with either technology.

The LTE specification defines a normal CP length of 5 µs, an extended CP of 17 µs intended for suburban and rural environments with typically longer delay spread, and a CP designed for MBSFN with a duration of 33 µs.

In order to maintain the slot duration constant at 0.5ms some adjustments to the CP duration are needed. In normal CP, the CP duration of the first symbol in the slot is slightly larger than the CP of the subsequent symbols, i.e., considering the symbol duration $T_u = 1/\Delta_f = 66.7$µs, the first CP in the slot is 5.2µs while the remaining are 4.7µs, so that $(5.2µs + T_u) + 6(4.7µs + T_u) \approx 0.5$ms. See Figure 3.4 on the facing page.

3.2.2. SYNCHRONIZATION AND CELL SEARCH

Synchronization is crucial for any digital communication system, see description of general OFDM synchronization on page 18. It is so important that 3GPP designed for LTE a specific procedure to indicate what steps need to be taken by the terminal in order to establish a successful registration in the network. In LTE the following two cell search procedures exist:
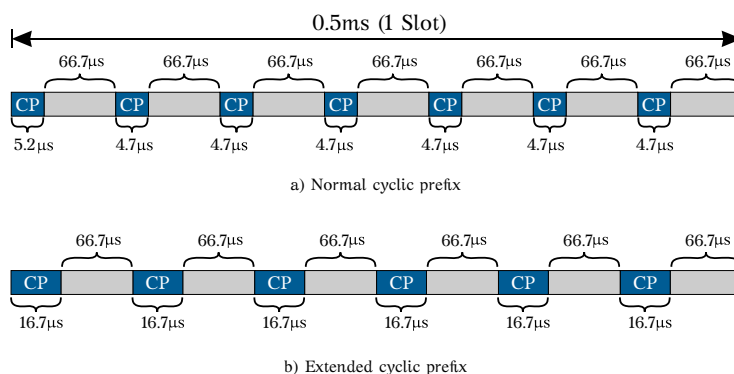
Figure 3.4.: LTE Symbol and CP sizes for: a) normal CP and b) extended CP.

INITIAL SYNCHRONIZATION, where the terminal scans the spectrum looking for a valid LTE base-station, finds it, and decodes the required information to register to the network. This is typically done when the terminal is switched on for the first time, or when it has lost connection to the network.

NEW CELL IDENTIFICATION, prepares the network for a possible handover between neighboring cells. As the terminals moves around the cell's edge it detects a neighboring BS, this action triggers the *new cell identification procedure* that instructs the terminal to inform the current connected BS the parameters related to the new BS. There are two physical signals used for these two aforementioned procedures, namely Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS), these are broadcasted by each cell, and provides the following information to the mobile terminal:

- timing of 0.5ms within each frame (PSS),

- physical layer Id (PSS),

- radio frame timing (SSS),

- cell Id (SSS),

- cyclic prefix length (SSS),

- TDD/FDD detection (SSS).

The PSS and SSS are transmitted twice per frame, while the exact transmission time depends on whether the system is TDD or FDD, e.g., if FDD they are transmitted as the last two symbols of slot 1 and slot 11, whereas for
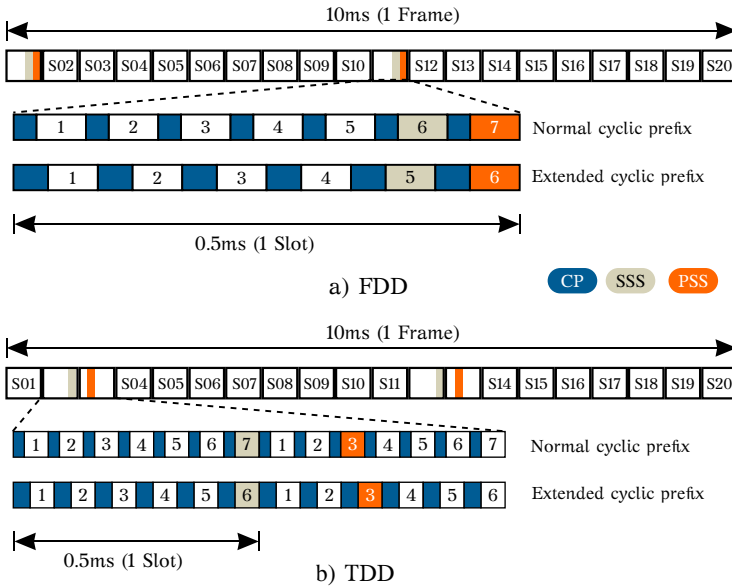
Figure 3.5.: Location of PSS and SSS in LTE for: a) FDD and b) TDD.

TDD they are transmitted in pairs in slots 2,3 and 12,13. An illustration of the exact position is provided in Figure 3.5.

The exact position of these synchronization signals also changes, and it does in accordance to the CP on duty. Recall that the number of symbols varies from 7 to 6 if extended CP is in use. Each synchronization signal occupies one OFDM symbol in the time domain, and spans the 72 central sub-carriers, in this way the same structure can be used independent with respect to transmission bandwidth in used. Remember that when a terminal is switched on for the first time, it has no previous knowledge what transmission bandwidth is on operation. Furthermore, out of those 72 sub-carriers, only 62 of them contain information, the remaining 5PREs at each extreme are left unused, so that a lower sample rate and a FFT with 64 points can be used to decode the synchronization signals and spare some energy in the mobile terminal.

3.2.3. CARRIER AGGREGATION

In LTE-A two or more carriers of LTE-R8/9 are aggregated in a variety of scenarios to a maximum bandwidth of 100 MHz. These scenarios are defined in accordance with the available bands and technologies currently deployed in various parts of the world.
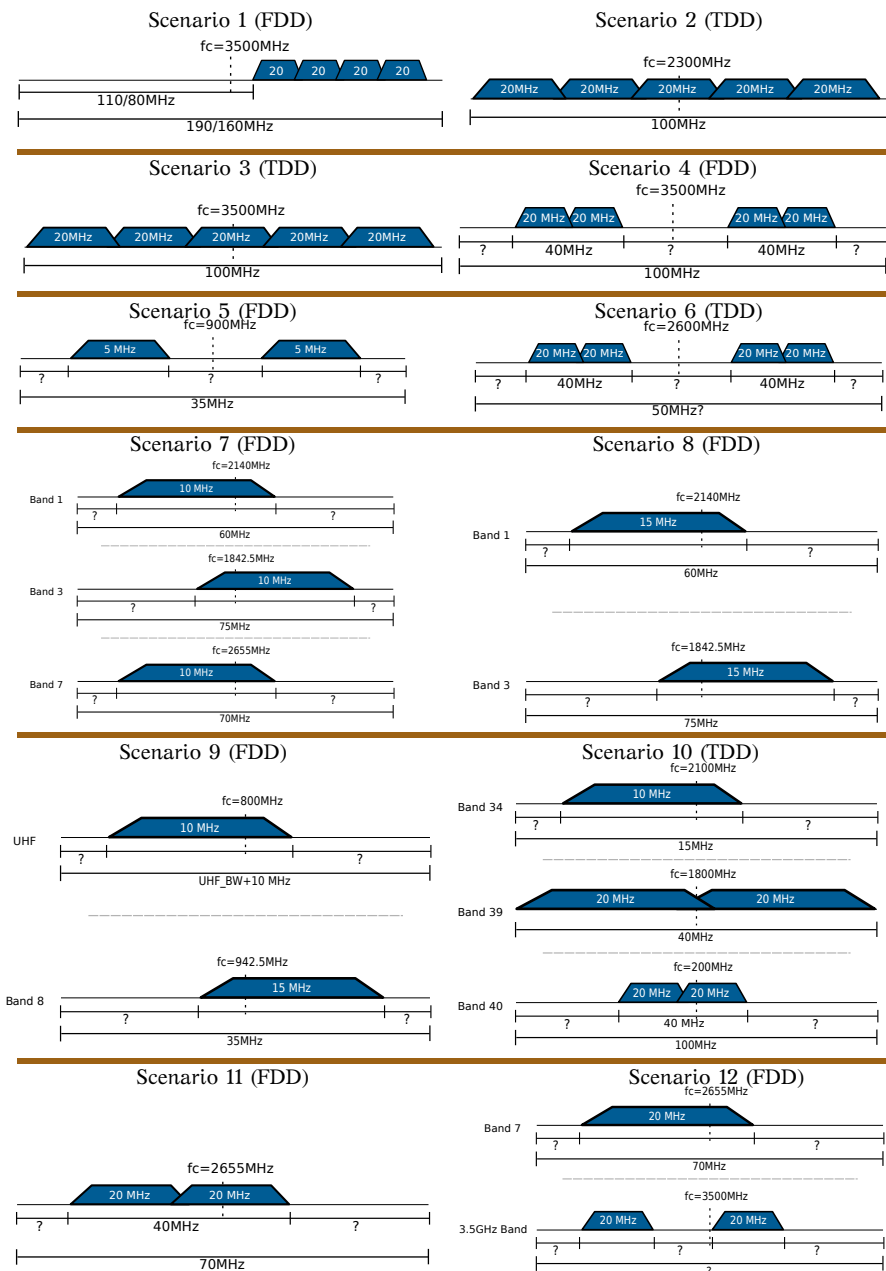
Figure 3.6.: Carrier aggregation initial scenarios in LTE-A [27].

Three different cases can be generalized out of the 12 initial scenarios proposed by 3GPP [27], and depicted in Figure 3.6 on the previous page. Note that in the figure some bandwidths and band-spacing are unknown denoted with a question mark. From these initial scenarios, it is possible to abstract three types of aggregation, namely, intra-band contiguous Carrier Component Aggregation (CCA), intra-band non-contiguous CCA, and extra-band non-contiguous CCA,

INTRA-BAND CONTIGUOUS CCA, is where two or more carriers from LTE-R8/9 are aggregated adjacent to each other. Each has a central frequency f_{ci}, where i denotes the Carrier Component (CC) index. The CCs are contiguous if $f_{ci} - f_{ci-1} \approx B_{CC}$, i.e., the distance between every two CCs is approximately equal to the component transmission bandwidth B_{CC}, common to all CCs. This scenario is possibly the least likely to be deployed in the starting stages, since today's radio spectrum is extremely crowded and 100 MHz are not easy to allocate.

INTRA-BAND NON-CONTIGUOUS CCA, is the scenario where various CCs are allocated within the same band, while not adjacent to each other. In Non-contiguous CCA $B_{CC} < f_{ci} - f_{ci-1} < B_{band}$, i.e., the distance between every two CCs is larger than the component transmission bandwidth but smaller than the transmission bandwidth available for the current transmission band. This scenario is more likely to appear in countries where network sharing is used or middle carriers are used by different services or technologies.

EXTRA-BAND NON-CONTIGUOUS CCA, is the scenario when $f_{ci} - f_{ci-1} > B_{band}$, i.e., the distance between any two CCs is larger than the current transmission bandwidth, and CCs are allocated in different transmission bands. In these scenarios the transmission bandwidth can be very large, making it unfeasible to receive the entire aggregated bandwidth with a single front-end. Thus, in this article the extra-band non-contiguous scenarios are simplified to multiple intra-band non-contiguous channels, i.e., there is one receiver front-end allocated to each band.

3.2.4. TERMINAL CLASSIFICATION

It is no surprise that given all technological advances introduced in LTE, a terminal capable of supporting all the features in the standard specification will undoubtedly suffer from high complexity. Thus, 3GPP created terminal categories, where complexity can be reduced by limiting functionality and terminal capabilities. This classification is shown in Table 3.1 on the facing page. In this manner, low-end and high-end terminal can all share the same infrastructure while consuming as little energy as possible given its

Table 3.1.: Terminal classification in LTE in releases 8, 9, and 10
extracted from [25].

Category	Release 8/9/10					Release 10 only		
	1	2	3	4	5	6	7	8
DL peak data-rate [Mbps]	10	50	100	150	300	300	300	3000
UL peak data-rate [Mbps]	5	25	50	50	75	50	150	1500
Maximum DL modulation [QAM]	64	64	64	64	64	64	64	64
Maximum UL modulation [QAM]	16	16	16	16	64	16	64	NR[2]
Max number of layers for DL spatial multiplexing	1	2	2	2	4	SS[1]	SS[1]	SS[1]

[1]Signaled separately. [2]Not reported in [25].

corresponding classification. An important consideration taken throughout this thesis is trying to simplify low-end terminals and thus reducing energy consumption.

3.2.5. CHANNEL MODELS

As previously discussed in Chapter 2, to use an appropriate model of the propagation channel is crucial to grasp the communication performance and its ultimate limits. The entire system performance can only be calculated accurately if channel models used for the performance are realistic.

The authenticity of the channel model relies on two factors, first the topology of the physical environment, i.e., indoor, outdoor, urban, suburban, rural, *etc.,* and second the frequency band where the signal is transmitted, as the radio waves interact differently with the environment depending on bandwidth and frequency.

The ITU developed channel models to be used for the third generation communication standard (3G) cover indoor office, outdoor-to-indoor, pedestrian, and vehicular scenarios. Later 3GPP adopted an extended set of channel models based on those initial ones from ITU but with an increased bandwidth to better model the 20MHz bandwidth used by LTE. This model uses mainly three channel models, which are fully described by its statistic characteristics such as Power Delay Profile (PDP), r.m.s. delay spread, and Doppler shift. A brief description of the models is given below:

- Extended Pedestrian A (EPA), describes a urban scene at slow pedestrian speeds; pedestrian speed considered to be 2 Km/h, equivalent to a Doppler shift of 5 Hz. Given the urban environment, this channel model contains a short delay spread, i.e., r.m.s. delay spread is defined as 43 ns, and a PDP that can be seen Table 3.2.

Table 3.2.: Power delay profile defined for EPA channel model [12].

Tap number	Excess tap delay [ns]	Relative power [dB]
1	0	0.0
2	30	-1.0
3	70	-2.0
4	80	-3.0
5	110	-8.0
6	190	-17.2
7	410	-20.8

- Extended Vehicular A (EVA), describes an urban scene at vehicular speeds; the vehicular speeds considered are 30 to 130 Km/h, equivalent to a Doppler shift of 70 and 300 Hz, respectively. This channel model contains a moderate long, or medium delay spread, its r.m.s. delay spread is defined as 357 ns, and a PDP that can be seen Table 3.3.

Table 3.3.: Power delay profile defined for EVA channel model [12].

Tap number	Excess tap delay [ns]	Relative power [dB]
1	0	0.0
2	30	-1.5
3	150	-1.4
4	310	-3.6
5	370	-0.6
6	710	-9.1
7	1090	-7.0
8	1730	-12.0
9	2510	-16.9

- Extended Typical Urban (ETU), despites its name it describes an extreme case of urban, suburban and rural areas. This model serves to define the performance limits in LTE implementations, this channel model is defined with the vehicular speed of 30Km/h, equivalent to a Doppler shift of 70 Hz. This channel model contains an very long delay

Table 3.4.: Power delay profile defined for ETU channel model [12].

Tap number	Excess tap delay [ns]	Relative power [dB]
1	0	-1.0
2	50	-1.0
3	120	-1.0
4	200	0.0
5	230	0.0
6	500	0.0
7	1600	-3.0
8	2300	-5.0
9	5000	-7.0

spread compared to the aforementioned models, i.e., the r.m.s. defined as 991 ns, and a PDP that can be seen Table 3.4.

3.3. IEEE 802.11 WIRELESS LAN

The 802.11 family encompasses a large number of members, not all compatible to one another, e.g., the first version of 802.11, released in 1997 was based on DSSS, and FHSS technologies. The original specification aimed for transmission data-rates of 2 Mbps but as the demand for higher data-rates arose, both DSSS, and FHSS were unable to provide the required performance and opened room for OFDM technology. Ironically, back in the 90's when the need for wireless connection to local networks first appeared, there were two competing standards being developed. Namely, HIgh PERformance Local Area Network (HIPERLAN) proposing multi-carrier transmission (OFDM) as the basis communication technology, and 802.11 proposing spread-spectrum techniques. Since the transmission band to be used was unlicensed and used by other RATs, the Federal Communications Commission (FCC) demanded spread-spectrum techniques to be used [9].

As the standard has evolved newer and more sophisticated technological enhancements have been added with the ultimate goal of increasing peak data-rates. Among these enhancements the following are included:

- Higher modulation order for improved spectral efficiency,

- additional bandwidth by using the 5 GHz band and more recently 60 GHz is also included in newer releases,

- spatial multiplexing to increase data-rate further with additional antennas,

- implementation of Quality of Service (QoS) to improve user experience,

Table 3.5 shows a summary of some important parameters; the table comprises releases up to 2014. Notice that the data rate has increased more than three orders of magnitude in only 15 years. As an example, the latest release, 802.11ac achieves a maximum data rate of almost 7 Gbps which is 3,500 larger than the original goal back in 1997. This has been feasible due to inclusion of modern techniques into the standard such as MIMO techniques, packet aggregation, 256QAM modulation, binary convolutional coding with coding rates of 3/4 and 5/6, in addition to a larger bandwidth and improved protocol procedures.

Table 3.5.: Comparison of 802.11 family standards up to 2014.

	Year of release	Band-width [MHz]	Freq. [GHz]	Radio technology used	Maximum data-rate [Mbps]
802.11	1997	20	2.4	DSSS, FHSS	2
802.11a	1999	20	5	OFDM	54
802.11b	1999	20	2.4	DSSS	11
802.11g	2003	20	2.4	OFDM,DSSS	54
802.11n	2009	20, 40	2.5, 5	OFDM	150
802.11ad	2012	2, 160	60	OFDM	6,912
802.11ac	2014	20, 40, 80, 160	5	OFDM	7,000

3.3.1. FRAME STRUCTURE

WiFi is designed to be an in-burst standard, i.e., the standard provides means to acquire system parameters from the BS[2] on a very early stage of the connection. The latest release, namely 802.11ac, provides a frame structure where a preamble is used for early, signal detection, synchronization, channel estimation, and acquisition of vital parameters such as modulation order, coding rate, *etc*. The preamble defined for Very High Throughput (VHT) is depicted in Figure 3.7 on the next page.

VHT is a term used to distinguish between different versions of WiFi. Given the capabilities of the terminal, 802.11ac, has to be backward compatible with 802.11n, which in turn is compatible with 802.11a. Thus, 802.11ac contains a preamble structure based on that of 802.11a, referred as legacy support. The frame structure of 802.11ac, consists of three type of fields,

---

[2]The term Base Station (BS) is typically reserved for cellular systems, in WiFi, the analogous device is commonly called access-point or wireless-router. However, for sake of simplicity, the term BS is used to describe both.

Figure 3.7.: Preamble for 802.11ac, [28].
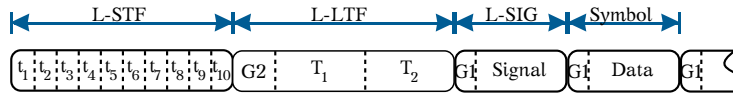
Figure 3.8.: Preamble for 802.11a, referred as legacy in subsequent releases [29].

a) The Short Training Field (STF), in charge of signal detection, Automatic Gain Control (AGC), coarse CFO estimation, and symbol timing estimation. It is called short because it consists of 10 symbols of short duration and the length of this field is of 8µs.

b) The Long Training Field (LTF), in charge of channel estimation and fine frequency offset estimation. It consists of 2 long symbols specially designed for channel estimation with a long guard interval, or cyclic prefix. The duration of this field is 8µs, equivalent to two regular OFDM symbols.

c) The signal field is used to communicate other system parameters, such as modulation order, coding rate, number of antennas, even the number of data symbols that follows. The duration of this signal field is of 4µs, the same as a regular OFDM symbol.

The original preamble from 802.11a (legacy) is depicted in Figure 3.8. Note that in the figure the Guard Interval (GI) is equivalent to the CP. There are two types of GI, GI 1 of 0.8µs duration and GI 2 of 1.6µs. GI 2 is longer in order to assure that the channel impulse response does not affect any part of the LTF, since this field is critical for channel estimation which is crucial for a successful communication. Notice that there is no explicitly GI defined for the STF, yet, since the STF consists of 10 identical symbols of 0.8µs duration each, this is equivalent to having the first short symbol used as CP, or even more short symbols if the channel impulse response is longer than 0.8µs. The STF has a special importance in this thesis, as it is the main resource used for synchronization and CFO estimation.

3.4. DIGITAL VIDEO BROADCAST

Digital Video Broadcasting (DVB) is a family of standards aimed at improving television transmission performance. The standard is regulated by the DVB Project, where the specifications can be downloaded at no cost. DVB targets transmission of television broadcast signals in a large variety of methods, e.g., satellite, cable, terrestrial, aerial, and microwave transmission.

The focus of this thesis is only on that part of the DVB family intended for mobile reception, i.e., in hand-held devices. Digital Video Broadcasting - Handheld (DVB-H) is a subset of the terrestrial specification of DVB, Even though Digital Video Broadcasting - Terrestrial (DVB-T) can provide service to fixed hand-held terminals and terminals moving at moderate speeds, the specification does not implement any power saving techniques, nor is capable of providing services at vehicular speeds. Moreover, DVB-H has to be able to efficiently switch from one BS, or transmitter cell to a new one as needed. DVB-H is also designed so that a single terminal supports multiple transmission bands in various parts of the world.

DVB-H uses the signal originally designed to be received by a DVB-T device, and appends additional techniques to be able to be received by a DVB-H terminal, these additional techniques can be listed as

- *time-slicing* technique consists of partitioning the transmission signal into smaller Internet Protocol (IP) datagrams. These datagrams are transmitted in burst-mode within small time duration. In this manner the receiver can fill-up a buffer and switch-off the listening part of the receiver, and thus save energy.

- An additional Reed-Solomon Forward Error Correction (FEC) is used to protect the transmitter signal from the wireless propagation channel not present in DVB-T.

- A special signaling and Transmitter Parameter Signalling (TPS) is defined to improve and speed-up service discovery, e.g., a cell Id is inserted in the transmitter signal to support quicker scan and frequency handover.

- DVB-H adds a new transmission mode based on a 4K point FFT in order to trade-ff mobility and cell size. In this manner reception in Single Frequency Networks (SFNs) at high speeds is possible.

Figure 3.9 on the next page illustrates a conceptual diagram of a DVB-H receiver, the DVB-H demodulator, the received signal is demodulated via a regular DVB-T demodulator, where the corresponding transmission mode is detected and its corresponding TPSs extracted. When the burst transmission is finished, the timing slicing module switches-off the DVB-T demodulator to
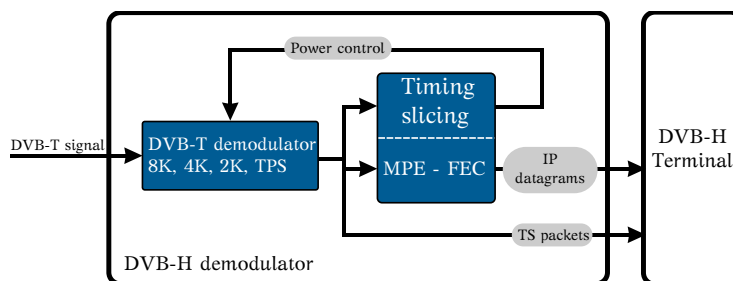
Figure 3.9.: Conceptual diagram of a DVB-H receiver [30].

save energy. Then, a large buffer inside the timing slicing module in combination with the FEC module and an interleaver feed the DVB-H terminal with IP-datagrams while the DVB-T part of the terminal is off.

3.4.1. PHYSICAL LAYER STRUCTURE

DVB-H provides a wide selection of parameters to adjust the transmission to the most suitable combination given a large set of environments of network topologies. DVB-H defines operation bandwidths of 5, 6, 7 and 8MHz and three operation modes i.e., 2K, 4K, or 8K. The transmission mode is independent of the transmission bandwidth, consequently, the symbol duration and sub-carrier spacing depends on both parameters. The standard specification defines a basic elementary period for each transmission bandwidth, namely, 7/40, 7/48, 7/56, and 7/64μs, for transmission bandwidths of 8, 7, 6, and 5MHz respectively. Then the symbol duration can be calculated by multiplying the elementary period with the transmission mode, so that all possible symbol durations are listed in Table 3.6.

Table 3.6.: Symbol duration and sub-carrier spacing given transmission mode and bandwidth [31].

	8K [μs]	4K [μs]	2K [μs]
5 MHz	1,433.6	733.6	358.4
6 MHz	1,194.67	611.33	298.67
7 MHz	1024	524	256
8 MHz	896	458.5	224

Note that the longest symbol is given by the combination of 5MHz bandwidth in 8K transmission mode, since the sub-carrier spacing is equal to the inverse of the symbol duration, this case also corresponds to the smallest possible sub-carrier spacing $\Delta_f = 697.54$Hz, which can be very sensitive to frequency variations.
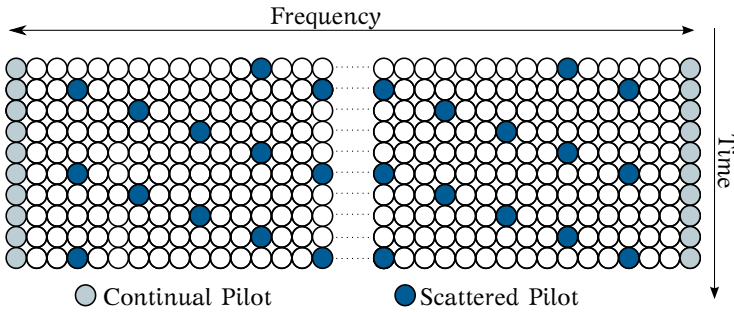
Figure 3.10.: Pilot positions in DVB-T and DVB-H time-frequency
grid [31].

Each OFDM symbol is preceded by a GI or CP whose duration depends
on the symbol duration and the GI type. The specification defines four types
of GI, i.e., 1/4, 1/8, 1/16, and 1/32 of the symbol duration, thus, based on the
number of possible symbol durations depicted in Table 3.6 on the previous
page, the total number of possible CP durations is 48.

In order to trade-off efficiency, additional degrees of freedom are provided
in the specification, e.g., the modulation rate supported varies from Quadra-
ture Phase Shift Keying (QPSK) to 64QAM. The coding rates of 1/2, 2/3, 3/4,
5/6, and 7/8 can also be used to combat difficult propagation occurrences.
According to [32], the configuration that is most likely to be used for mobile
reception consists of a constellation order of 16 QAM, in some cases with
good SNR the maximum of 64 QAM can be used, although it is not recom-
mended. The FECs best suited for mobile communication is 1/2 and 2/3 rates.
Finally, depending on the network topology, the recommended configuration
for mobile service is 2K mode with 1/4 of GI, or 4K mode with 1/4, or 1/8
GI, or 8K mode 1/4, or 1/8 GI.

The transmitted signal is organized in frames of 68 OFDM symbols, 4
frames constitute a super-frame. In addition to the transmitted data, a
DVB-H frame contains: a) TPS carriers, b) continual, and c) scattered pilots.
The pilots can be used for synchronization, including frequency and channel
estimation. An illustration of the time-frequency representation in DVB-H is
given by Figure 3.10. The continual pilots are used for frequency locking and
fine frequency estimation, while the scattered pilots are intended for channel
estimation.

The TPS are located within the first 68 consecutive OFDM symbols (first
frame) of each super-frame. They carry information related to a) modulation
order, b) hierarchy information, c) guard interval, d) code rate, e) trans-
mission mode, f) frame number, and g) cell identifier. Even though it is
not specified in the standard, blind methods to estimate symbol timing and
other parameters can also be used, since the standard specification typically

described in detail the transmission procedure, whereas the receiver is usually left up to consideration of the designer, as long as it complies with the standard specification. This allows the different terminal vendors to play with design parameters and be attractive to the market in terms of price and functionality.

3.5. OFDM STANDARD SPECIFIC PARAMETERS

The selection of specific parameters for each standard follows a delicate process under which factors like, data-rate, mobility, and even intended environment play an extremely important role. One would want to have the symbol duration as long as possible, in order to reduce the overhead imposed by the addition of CP. This overhead is given by

$$\frac{T_{CP}}{T_u} = \frac{T_{CP}}{T_{CP} + T_s} = \frac{T_{CP}}{T_{CP}\Delta_f + 1}\Delta_f, \tag{3.1}$$

where T_{CP} is the duration of cyclic prefix, Δ_f is the sub-carrier spacing, and T_s and T_u are the duration of OFDM symbol with and without CP, respectively. Since $T_u = 1/\Delta_f$, it is desired that the sub-carrier spacing (Δ_f) to be as small as possible, but not too small, since it also contributes to increase sensitivity to frequency errors, including the one introduced by Doppler shift.

Another manner to reduce the CP overhead, is to reduce the duration of the CP. However, as seen previously in Subsection 2.2.2 on page 16, in order to maintain orthogonality, the CP has to be larger than the channel's impulse response. The duration of the channel's impulse response depends on the geography of the communication environment, which is very difficult to predict, so worst case scenarios are typically used based on channel statistics of the intended transmissions environments (indoor, outdoor, urban area, open field, *etc.*,).

The CP overhead results in additional power and reduced bandwidth efficiency. With the power loss in mind, if the cell size increases the system becomes more power limited, since more power is needed to reach the cell's border. Then there is trade-off between power loss due to CP and the interference caused by the part of the channel impulse response not covered by the CP. Typically, the larger the cell area the longer the channel's impulse response. Therefore, beyond a certain cell size there is no reason to increase the CP further as the power loss would have a larger negative impact [25]. As a consequence, RATs specification, e.g., LTE and DVB-H, often define several lengths of CP to be used in accordance to transmission environment.

Once the sun-carrier spacing and the CP has been selected based on the environment, expected Doppler spread, *etc.*, the total number of sub-carriers is determined conditional to the amount of bandwidth available for trans-

Table 3.7.: OFDM Parameter specification per standard.

Name	Symbol	LTE	WiFi ac	DVB-H
Sub-carrier spacing [KHz]	Δ_f	15	312.5	$1/T_u$x10^3
Symbol duration [µs]	T_u	66.667	3.2	N/F_s
Cyclic Prefix duration [µs]	T_{CP}	5, 17*, 33†	0.4, 0.8, 1.6	$T_u/32$, $T_u/16$, $T_u/8$, $T_u/4$
Transmission bandwidth [MHz]	B	1.4, 3, 5, 10, 15, 20	20, 40, 80	5, 6, 7, 8
Sampling frequency [MHz]	F_s	30.72‡	20, 40, 80	40/7, 48/7, 56/7, 64/7
FFT size	N	2048‡	64, 128, 256	2048, 4096, 8192

*Extended CP for suburban and rural areas.
†For broadcasting services only.
‡Defined for 20MHz transmission bandwidth, however, it can be reduced in smaller bandwidths.

mission. The total bandwidth is given by

$$B = N\Delta_f = N/T_u, \tag{3.2}$$

with N as the number of points of the FFT used for OFDM modulation. It is important to note that spectrum occupied by an OFDM transmission falls slowly outside the band of interest, leading to problematic *out-of-band emission* that produces interference in neighboring services. Thus, in practice the transmitted signal needs to be filtered to suppress *out-of-band emissions*, and in some standards like LTE the sub-carriers at the borders are left unused. The sample-rate is directly related to the bandwidth of the transmitted signal, and it follows that a large bandwidth increases sample-rate and thus, increases terminal complexity. Even more so if an oversampling is considered; a common practice in actual implementations. Often, a large variety of FFT sizes is provided in order to fit the available bandwidth.

These OFDM specific parameters per standard are listed in Table 3.7.

Notice that the requirements set by the three standards listed varies broadly, especially in DVB-H, where depending on the bandwidth available and the mode, the sample rate changes in non-multiple of one another. The largest sampling rate is imposed by WiFi in its ac version, where in order to process the largest transmission bandwidth; a sampling rate up to 80MHz needs to be supported.

The FFT size and sampling rate of LTE is only defined on its largest transmission bandwidth. However, in practice, these two parameters can be changed by mobile terminal manufactures as they feel convenient in terms of cost and performance, e.g., on a transmission channel with 5MHz transmission bandwidth it is enough to use an FFT with 512 points yielding to a sampling frequency $F_s = 7.68$MHz, instead of the 30.72MHz defined in the official specification.

Part I.

Sign-bit estimation

This part presents a new approach to estimate various parameters based on the Sign-Bit (SB) of the received signal. This methodology combines algorithm development, thorough analysis and implementation in order to judge its efficacy in terms of complexity and performance. An ASIC is being fabricated as a proof of concept in 65 nm technology. The prototype is capable of estimating symbol timing, CFO and SNR based on the SB, occupies an area of $0.03\,\text{mm}^2$. Post-layout simulations with toggling annotated information report a power consumption as of 479 μW average power for LTE-R8/10 at 1.2 volts supply voltage. Furthermore, Performance and complexity analysis is presented in comparison to an equivalent 8-Bit quantization methodology. All the content in this part is a compilation from the following published or submitted articles:

- DIAZ, I., WILHELMSSON, L., SOFOTASIOS, P., MIAO, Y., TAN, S., EDFORS, O., AND ÖWALL, V. "A new approach to Sign-Bit based parameter estimation in OFDM receivers," *submitted to Springer Circuits, Systems, and Signal Processing on March 2014*.

- WILHELMSSON, L., DIAZ, I., OLSSON, T., AND ÖWALL, V., "Analysis of a novel low complex SNR estimation technique for OFDM systems," *IEEE Wireless Communications and Networking Conference*, March, 2011.

- DIAZ, I., WILHELMSSON, L., RODRIGUES, J., LÖFGREN, J., OLSSON, T., AND ÖWALL, V., "A sign-bit auto-correlation architecture for fractional frequency offset estimation in OFDM," *IEEE International Symposium on Circuits and Systems*, May, 2010.

- DIAZ, I., WILHELMSSON, L., RODRIGUES, J., OLSSON, T., AND ÖWALL, V., "Sign-Bit based architecture for OFDM acquisition for multiple-standards," *NORCHIP*, November, 2009.

4

A type of low-complexity estimation

Mobile terminals are in general subject to demanding constraints related to area, power and throughput. Although architectures and algorithms typically employ fixed-point arithmetic to address those stringent requirements, determining the optimum word-length that yields maximum performance at minimum cost is remarkably cumbersome since it constitutes a non-deterministic polynomial-time hard (NP-Hard) problem [33]. The procedure of selecting the best word-length typically involves the definition of an error boundary (maximum acceptable error), a cost function, and estimation of which word-length meets requirements within the error boundary either, heuristically [34], or analytically [35] [36].

OFDM is an emerging technology that is regarded as the standard technique for high data-rates and high-mobility applications, particularly in handheld devices driven by batteries. This has led to an increased industrial and academic interest towards finding effective and robust methods for improving the energy consumption in OFDM receivers. Motivated by the above observation, the present work is based on a new approach in the word-length selection process, which embraces extreme quantization and acts upon the error boundaries by systematic manipulation of the estimation algorithm. As a result, the corresponding performance degradation is controlled and traded-off with minimum area and reduced power dissipation. It is noted here that the proposed algorithms are not a straight-forward implementation of the more traditional higher precision counter-parts. Indeed, the proposed Sign-Bit (SB) estimation techniques are modified accordingly in order to assure its appropriate functionality under realistic scenarios. These modifications are novel and are backed-up with a corresponding mathematical analysis which is validated through comparisons with results from computer simulations.

More specifically, SB estimation is analyzed in the context of low-complexity OFDM receivers where its potential is demonstrated by performance com-

parison and post-layout functional simulations. To this end, three critical parameters in establishing efficient and reliable communication, namely, symbol-timing, CFO, and SNR are estimated with the aid of SB only. Note that in a real terminal a larger set of parameters or non-idealities are present, which are described and modeled in detail in [37]. However, SB estimation and its potential can be easily demonstrated for symbol-timing, CFO, and SNR.

It is recalled here that CFO and SNR estimators can be classified based on the type of data used for estimation (data-aided or blind), and the domain where the estimation takes place (time or frequency). The main drawback of frequency-domain estimators is higher complexity when compared to those operating in time-domain. This is evident in [38] and [39], where a cost function based on frequency-domain interference is used for CFO estimation. On the contrary, time-domain estimators are typically simpler and they can estimate symbol timing in addition to CFO such as in [20], [17] and [21]. Considering the above, the methods presented in this work belong to the classification of time-domain blind estimators as time-domain guarantees low complexity and blind provides a generality to the estimation covering all existing OFDM standards.

4.1. APPLICABILITY IN OFDM SYSTEMS

Parts of a typical receiver are depicted in Figure 4.1 on the next page where in each front-end the signal from one antenna is amplified, filtered, down-converted and digitized. Subsequently, symbol timing and CFO are estimated and compensated with the initial process being typically referred to as time-frequency synchronization during acquisition [16]. Efficient and robust synchronization is typically assisted by the use of pilot symbols/tones or a preamble. This has led to the design of numerous techniques based on pilots, see [40], [17], [41], [42], and the references therein. However, the present work refers to a wider range of emerging communication systems that employ a CP, which has been adopted by all OFDM standards.

The IFFT modulation function of an OFDM system is defined as,

$$F_k(t) = \begin{cases} \frac{1}{\sqrt{T_u}} e^{-j2\pi\Delta_f k(t-T_{CP})} & \text{if } t \in [0, T_s] \\ 0 & \text{otherwise} \end{cases}, \tag{4.1}$$

where T_s and T_u represents the symbol duration with and without CP, respectively. Δ_f is the sub-carrier spacing, and T_{CP} is the duration of CP. Recall from Subsection 2.2.2 on page 16 that the CP is a copy of the last part of the symbol, thus $F_k(t) = F_k(t + T_u)$, when $t \in [0, T_{CP}]$, i.e., inside the CP. Then, it follows that the complex baseband signal of the lth OFDM symbol is
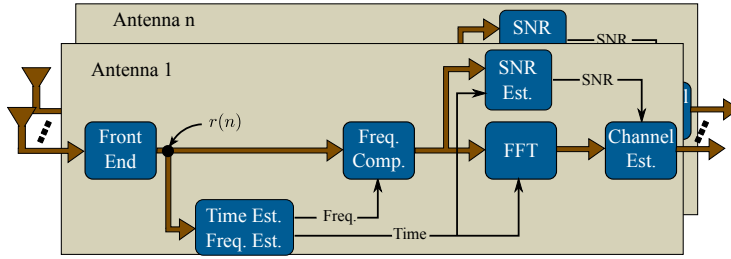
Figure 4.1.: A generic OFDM system.

represented as,

$$s_l(t) = \sum_{k=0}^{N-1} \alpha_{k,l} F_k(t - lT_s),$$ (4.2)

where $\alpha_{l,k}$ denotes the modulated data sub-carrier k of the lth symbol and N is the total number of sub-carriers or FFT size. The representation of the baseband signal with infinite duration, and thus, with infinite number of OFDM symbols is described as,

$$s(t) = \sum_{l=-\infty}^{\infty} s_l(t) = \sum_{l=-\infty}^{\infty} \sum_{k=0}^{N-1} \alpha_{k,l} F_k(t - lT)$$ (4.3)

The corresponding received signal in time-domain, directly after the front-end as shown in Figure 4.1, can be expressed as follows,

$$r(t) = h(t - \tau) * s(t) e^{j2\pi\Delta\varphi t} + w(t),$$ (4.4)

where h is the channel impulse response, τ is the time difference in between the transmitted and the received signal, $*$ denotes convolution, note that the channel impulse response as defined in Subsection 2.1.7 on page 12 is time-variant, whereas the channel above is not. For sake of simplicity, the initial formulation is done with an invariant channel, even though the posterior validation is presented with more complex channels. $\Delta\varphi$ is the corresponding CFO, and w is a zero-mean complex Additive White Gaussian Noise (AWGN) with variance σ_w^2.

The CFO is normalized with respect to the sub-carrier spacing (Δ_f) and its value is assumed to span over the range[1] $-0.5 < \Delta\varphi \le 0.5$. Furthermore, although the energy of the noise can be estimated in frequency domain, the SNR estimator is placed, in time-domain i.e., prior to the FFT. This is shown in Figure 4.1 on the previous page and is particularly useful as it exhibits an early indication of the current SNR which enables the post-FFT stages to adapt quickly to the corresponding transmission conditions.

The analysis presented in the current study applies to any OFDM system. However, in order to evaluate the proposed methods in a realistic manner, the LTE communication standard is chosen as the target application. LTE has been developed by the 3GPP in its release 8 and defines a basic sample rate of 30.72 MHz for the maximum component carrier bandwidth of 20 MHz with a sub-carrier spacing of 15 kHz [43]. The size of the FFT is 2048. In LTE there is a normal and an extended CP with durations of 4.7 µs and 17 µs, respectively (equivalent to 144 and 522 samples). The normal CP is the most commonly used mode and thus, selected in the present analysis, a more detailed description of LTE RAT can be consulted in Section 3.2 on page 29. Performance evaluation is carried out by Matlab simulations with channel characteristics described in more detail by [44]. These channels are simulated under 20 MHz bandwidth subject to Rayleigh fading. It is recalled that 3GPP provides three reference channel models of typical multi-path propagation environments summarized in Table 4.1. Furthermore, the proposed analyzes are compared to a more conventional yet equal estimation where the signal is quantized to 8-bits. In what follows, the same 8-bit methods are also used for corresponding complexity analysis and comparisons.

Table 4.1.: 3GPP channel models with its corresponding Doppler shift and delay spread [44].

Channel Model	Maximum Doppler Shift	r.m.s. Delay Spread
Extended Pedestrian A (EPA)	5 Hz	43 ns
Extended Vehicular A (EVA)	5 Hz	357 ns
Extended Typical Urban (ETU)	70 Hz	991 ns

---

[1]The frequency offset $\Delta\varphi$ may be also larger and can be divided into an integer and a fractional part. However, in the presence of an integer part the orthogonality between carriers is still preserved. To this effect, the present analysis accounts only for the fractional part and assumes that the integer part is zero.

5

Sign-bit symbol timing and carrier frequency offset

In this chapter the algorithm derivation of symbol timing and CFO estimation are presented and scrutinized in terms of performance, for comparison purposes, the full precision equivalent algorithm is also derived and compared to that of SB.

Since the algorithms apply to the digital part of the baseband receiver, the signals used in this context are assumed to be discrete and quantized, Therefore, the correspondent translation of time duration into number of samples is needed, to this extend, the duration of CP and a single symbol is given by

$$L = \lceil T_{CP} F_s \rceil \,, \tag{5.1}$$

and

$$N = T_u F_s. \tag{5.2}$$

Note that, due to the orthogonality requirement ($T_u = 1/\Delta_f$), the number of samples in the OFDM symbol equals the number of points in the FFT, and thus, N is used to represent either. Furthermore, the first sample of each symbol, immediately after CP is represented by θ, while the last sample of the lth symbol is defined as

$$\zeta_l = \theta_l + N - 1. \tag{5.3}$$

Based on these basic definitions, the full and SB estimation techniques can be derived in the following sections

5.1. FULL PRECISION ESTIMATION

A ML estimator using CP was proposed in [21] where the symbol timing (θ) and CFO ($\Delta\varphi$) estimates are determined based on the correlation between the CP and its counterpart in the OFDM symbol. The CFO is typically estimated more adequately in the time domain i.e., prior to the FFT operation [16]. An alternative, low-complexity method that is based only on maximum correlation was presented in [45] and is given by,

$$\hat{\zeta} = \arg\max_{n}\{|\gamma(n)|\},\tag{5.4}$$

where

$$\gamma(n) = \sum_{k=n-L+1}^{n} \beta(k),\tag{5.5}$$

and

$$\beta(k) = r(k)\overline{r(k-N)}.\tag{5.6}$$

The over-line notation in (5.6) denotes the complex conjugation. Note that in fact (5.4) estimates the last sample in the OFDM symbol, given that the maximum correlation is typically found just after the full symbol has been processed. Based on this, it follows that the actual symbol start can be calculated by,

$$\hat{\theta}_{l} = \hat{\zeta} - N + 1,\tag{5.7}$$

or

$$\hat{\theta}_{l+1} = \hat{\zeta} + L + 1,\tag{5.8}$$

depending on whether the symbol timing is estimated in relation to the currently processed symbol, as in (5.7), or to the incoming symbol, as in (5.8). Since the real start of a symbol needs $\hat{\zeta}$ as reference, it can be straightforwardly claimed that symbol timing has been found once $\hat{\zeta}$ is estimated.

The CFO can be computed in two manners, namely, the peak-based frequency estimator and the angle-based frequency estimator. The former, which is considered in the present work, extracts the phase of the correlation peak vector and the estimator can be expressed as,

$$\widehat{\Delta\varphi}_{p} = \frac{1}{2\pi}\arg\gamma(\hat{\zeta}).\tag{5.9}$$

The latter extracts the average argument of the samples $\beta(k)$ with k placed in the CP. Thus, the CFO estimator becomes,

$$\widehat{\Delta\varphi}_a = \frac{1}{2\pi L} \sum_{k=\hat{\zeta}-L+1}^{\hat{\zeta}} \arg\beta(k). \tag{5.10}$$

The estimators in the current section are designed assuming infinite precision in the signal representation and they can be considered as conventional estimators. However, it will be shown in coming sections that SB estimators are capable of maintaining performance degradation at reasonable levels with a significantly reduced complexity compared to the conventional estimators.

5.2. SIGN-BIT DERIVATION

As already mentioned in Section 4.1 on page 50, the correlation peak is used to determine the symbol timing. However, the considered signal in this case is now quantized to only the SB and thus (5.4), (5.5), and (5.6) can be expressed as,

$$\hat{\zeta}^q = \arg\max_n\{|\gamma^q(n)|\}, \tag{5.11}$$

$$\gamma^q(n) = \sum_{k=n-L+1}^{n} \beta^q(k), \tag{5.12}$$

and

$$\beta^q(k) = r^q(k)\overline{r^q(k-N)}, \tag{5.13}$$

where superscript q denotes that the received samples are quantized to the SB only. Consequently, the received sample takes only values $r^q(n) \in \{\pm 1 \pm i\}$, while the vector resulting from the complex multiplication in (5.13) has a phase with resolution of $\pi/2$, i.e., β^q takes only the values 0, $\pi/2$, π, and $3\pi/2$.

Analogous to full-precision, the frequency offset can be estimated in two ways subject to considerable advantages and disadvantages regarding the involved performance and complexity. In *peak-based frequency estimator* $(\widehat{\Delta\varphi}_p^q)$, the estimate is extracted from the phase of $\gamma^q(\hat{\zeta}^q)$ yielding,

$$\widehat{\Delta\varphi}_p^q = \frac{1}{2\pi} \arg\gamma^q(\hat{\zeta}^q), \tag{5.14}$$

where as before, $\hat{\zeta}^q$ is the estimated symbol timing and γ^q is the output of the correlator in (5.12). By initially assuming a non-dispersive noiseless
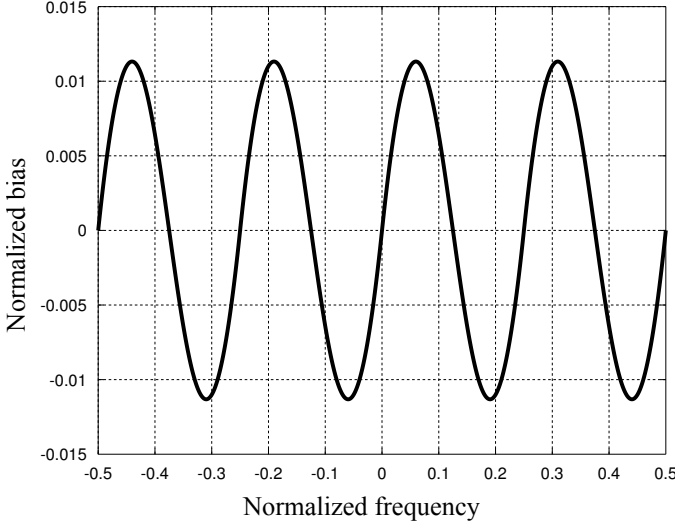
Figure 5.1.: Expected bias normalized to the sub-carrier spacing, under noiseless conditions with non-dispersive channel.

scenario, i.e., $h(n) = \delta(n)$ and $w = 0$, the bias[1] is illustrated in Figure 5.1, where it is shown that its amplitude depends on the true fractional CFO. This bias is periodic in intervals 1/4 of sub-carrier spacing, i.e., it is periodic in each quadrant of the complex plane. For the sake of simplicity, the analytical expression is derived for the first quadrant only ($0 \leq \Delta\varphi < 1/4$), namely, [46],

$$\mathsf{E}\left[\Delta\varphi_{\mathrm{p}}^{\mathrm{q}}\right] = \arctan\left[\frac{\Delta\varphi}{\pi/2 - \Delta\varphi}\right]. \tag{5.15}$$

Each periodic interval possesses two maximum biases, one positive and one negative, which can be analytically deduced by (5.15). The maximum bias is at $\Delta\varphi = \left(\pi \pm \sqrt{-\pi^2 + 4\pi}\right)/8\pi \approx \{0.06, 0.19\}$ and corresponds to approximately $0.011 \approx 1\%$ of the normalized frequency in the noiseless scenario. It is noted here that although the bias is considered to affect the system performance, it is shown in Section 5.3 on the facing page, that under realistic communication scenarios, the loss incurred by the bias is practically negligible.

---

[1]Bias is defined as the difference between the estimator's expected value and the true value.

The second method of estimating the CFO, *Angle-based frequency estimator* $(\widehat{\Delta\varphi}_a^q)$, consists of taking the phase of the individual samples within the CP. To this effect, the estimated value of $\Delta\varphi_a^q$ is computed as follows,

$$\widehat{\Delta\varphi_a^q} = \frac{1}{A} \sum_A \arg \beta_{\in CP}^q, \tag{5.16}$$

where A is the number of samples used for the estimation. In most estimators, a large value of A improves estimation reliability by decreasing the estimation variance. However, this comes at a cost of longer latency as several symbols are used to obtain a single estimate. As a consequence, the processing time is increased.

In contrast to the peak-based SB estimator, the angle-based estimator is unbiased. This can be witnessed by assuming a noiseless environment with no multi-path components and considering, for simplicity, only the first quadrant of the complex plane i.e. $0 \leq \Delta\varphi < 1/4$. Under this assumption the argument of β^q can only be 0 or $\pi/2$ whereas the probability that $r^q(n)$ and $r^q(n-N)$ have different phase is given by,

$$P\{\arg \beta^q(k) = \pi/2\} = \frac{\Delta\varphi}{\pi/2}. \tag{5.17}$$

Based on this, the expected value of the estimator can be expressed as follows,

$$E[\widehat{\Delta\varphi_a^q}] = \frac{\pi}{2} P\{\arg \beta^q(k) = \pi/2\} = \Delta\varphi. \tag{5.18}$$

At first glance the angle-based $(\widehat{\Delta\varphi^q}_a)$ seems a better alternative since it is unbiased. Nevertheless, it is extensively shown in following section that its performance is affected in a larger extent by the delay spread, in contrast to the case for peak-based $(\widehat{\Delta\varphi^q}_p)$.

5.3. ESTIMATION PERFORMANCE

In OFDM systems, time and frequency estimation are typically considered as a joint estimation process given that frequency estimation cannot be performed prior to time estimation. However, in this section they are presented separately in order to analyze their performance individually.

It was decided to use LTE as the performance test-bench for two reasons, first, because the author believe the major benefits of SB estimations are within low-power terminals, and LTE standardization has shown some interest in defining terminals under this category, and second, because the channel models provided by 3GPP provide a realistic scenarios worth exploring.
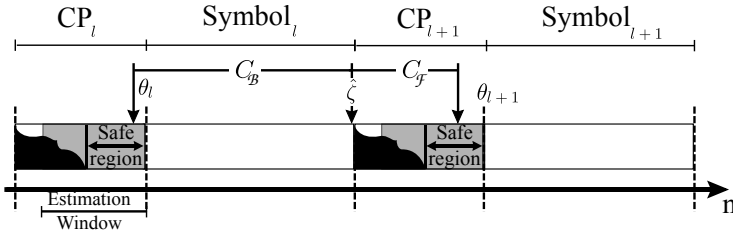
Figure 5.2.: Graphical representation of the safe region, the estimation window, and their relation to the channel impulse response.

SIGN-BIT TIME ESTIMATION

It is recalled that estimating the exact symbol start is not particularly crucial in OFDM systems. In order to avoid ISI, it is sufficient if the estimated symbol start lies anywhere within the end of the channel impulse response and the end of the CP, which is denoted as safe region in Figure 5.2. However, some performance analysis can be carried out based on the statistics of the estimator. To this end, it is firstly assumed that a symbol l and its corresponding CP are used for the time estimation; since the symbol plus the CP have a total of $N + L$ samples, the symbol timing is expected to be found at position $\hat{\zeta} = \theta_l + N - 1$ (see Figure 5.2). Consequently, a positive/negative constant (C_F or C_B) can be safely added in order to reposition the $\hat{\theta}$ closer to the center of the safe region, allowing timing errors in both directions. The estimation window is the region where samples are taken from in order to perform the corresponding estimation. The ideal case with the most reliable estimation would be if the estimation window covers only the safe region (See Figure 5.2).

However, in the presence of multi-path components, the length of the safe region varies with each channel realization while the estimation window remains constant. Thus, there will be "unclean" samples in the estimation, i.e., those samples affected by the channel impulse response. As a consequence, the channel introduces an offset which is dependent upon the r.m.s delay spread. Figure 5.3 on the facing page shows the average estimated symbol timing ($\hat{\zeta}$) for the EPA, EVA, and ETU. The symbol-timing appears as a positive shift from the expected symbol timing when no delay spread is present, i.e., in the case of a Rayleigh flat fading channel.

Observing any of the curves for one channel model in Figure 5.3 on the next page, it is noticed that the difference between SB and 8-bit average symbol-timing is practically zero for SNR higher than $4\,\text{dB}$, i.e., only a fraction of a sample. However, this is not the case for SNR lower than $4\,\text{dB}$ due to the higher quantization in the SB method.
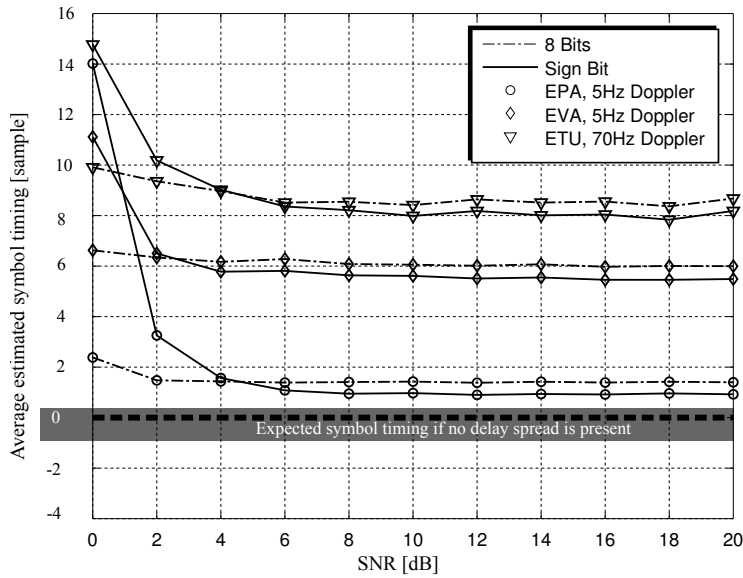
Figure 5.3.: Estimated Symbol timing for EPA, EVA and ETU chan-
nel models, relative to that expected symbol timing un-
der none multi-path environment.

This is demonstrated more clearly in the figure at 0 dB SNR, where the
quantization noise dominates the estimator's statistics. Based on this, the
three channel models have similar average estimated value, regardless of the
different value of r.m.s delay spread. Moreover, the distance from the expected
to the estimated value is quite small compared to 144, which is the length of
CP for LTE. The corresponding standard deviation is shown in Figure 5.4 on
the following page and, as expected, it decreases as the SNR increases, with
minimum deviation of 2 samples above 16 dB SNR. It is also noted that the
cross-over points mark the limit where SB matches or slightly out-performs
the 8-bit precision in terms of standard deviation.

In the high SNR regime, the SB appears to have better performance than the
8-bit counterpart. This is because the sign-bit quantization removes all small
variations from the received signal -the noise amplitude is much smaller than
the transmitted signal amplitude- resulting in a more distinctive correlation
peak i.e. smaller variance, for SB in comparison to 8-bit. However, in
the low SNR regime the opposite holds since the quantization removes a
large portion of the signal in the SB method. It is also expected that the
standard deviation increases with the channel's r.m.s delay spread, since in a
multi-path environment the channel impulse response affects samples in the
estimation window.
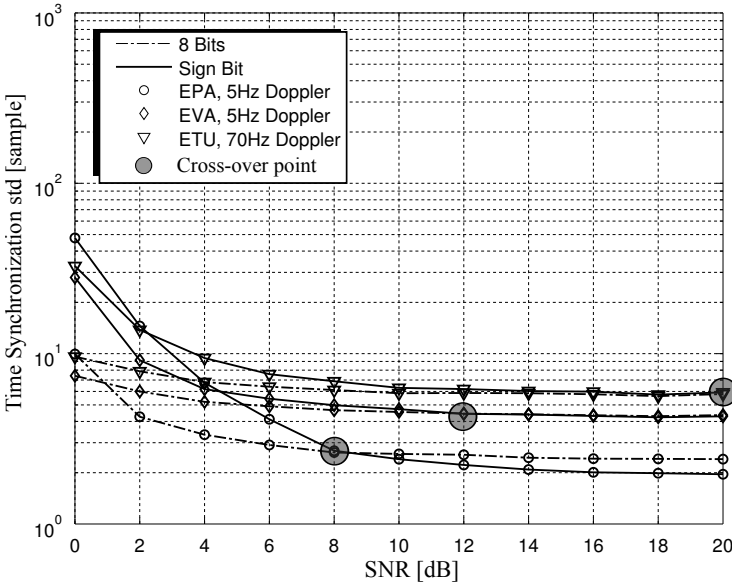
Figure 5.4.: Symbol timing standard deviation for EPA, EVA and ETU channel models as a function of SNR.

The correlation function $\gamma^q(n)$ in (5.12) is not capable to disregard samples affected by the channel, hence, the larger the delay spread, the more samples are affected and thus, the larger the standard deviation. This appears to affect both SB and 8-bit methods with the difference in standard deviation for the various channel models.

SIGN-BIT FREQUENCY ESTIMATION

CFO appears at the receiver as a consequence of frequency mismatch at the local oscillators and Doppler shift due to relative motion between transmitted and receiver. Given the channel models briefly introduced in Section 4.1 on page 50, the maximum Doppler shift considered is 70 Hz, which is considerably small compared to Δ_f, and does not represent a hindrance to the system. However, the maximum CFO due to oscillator mismatch is much larger and needs appropriate compensation, i.e., it equals $53.75\,\mathrm{KHz} \approx 3.5\Delta_f$[2]. Also recall from Section 4.1 on page 50 that this CFO is divided in an integer and a fractional part, where the presented estimator operates over the fractional part only, whereas the integer part is compensated at subsequent stages of the receiver.

---

[2]Considering a typical crystal accuracy of $\pm 20\,\mathrm{ppm}$ [37], and $5\,\mathrm{MHz}$ transmission bandwidth at the edge of the band 41 available in USA [25].
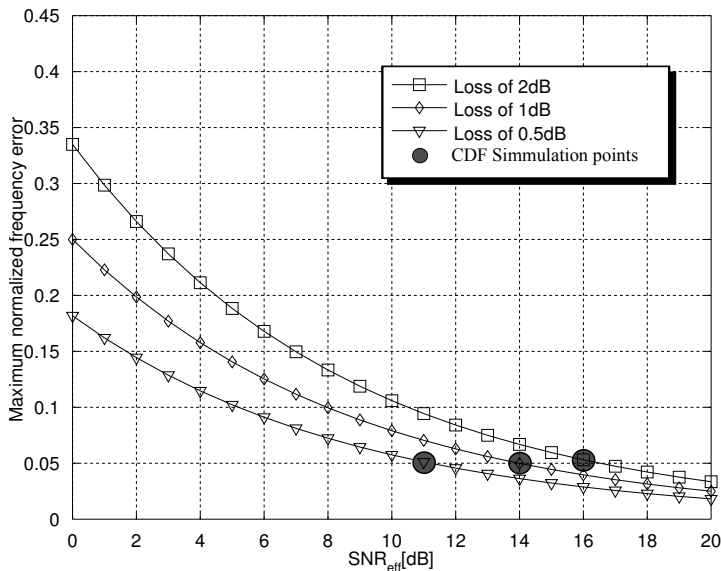
Figure 5.5.: Maximum allowed fractional frequency error normalized to the sub-carrier spacing ($\Delta_f = 15\,\text{kHz}$) for a allowed loss of: from top to bottom 2dB, 1dB, and 0.5dB
due to ICI interference.

As phase information of each sample is lost due to quantization, it is
reasonable to expect that SB frequency offset estimation would perform worse
than higher precision. Nevertheless, it is more relevant to determine whether
the corresponding performance loss is significant or not.

After estimating the CFO, the received samples have to be compensated in
order to restore orthogonality in the frequency domain. If the estimation is
not accurate, the compensation is not able to completely eliminate the CFO
and the impact of the remaining offset can be measured by the introduced
ICI. For a low CFO relative to the sub-carrier spacing ($\Delta_f = 15\,\text{kHz}$), the ICI
power can be approximated as [47]

$$I = \frac{\pi^2}{3}(\Delta\varphi)^2. \tag{5.19}$$

By treating I as additional noise, the effective SNR becomes

$$\text{SNR}_{\text{eff}} = \frac{S}{\eta + I}, \tag{5.20}$$

where η is the noise introduced by the channel and other imperfections, and S is the power of the desired signal. Figure 5.5 on the previous page illustrates the maximum allowed fractional frequency offset, normalized to sub-carrier spacing, so that there is a respective maximum SNR loss of $2\,\mathrm{dB}$, $1\,\mathrm{dB}$ and $0.5\,\mathrm{dB}$ due to the remaining frequency offset in the system. To measure the performance of the estimator, take for example, a maximum allowed frequency error of 0.05 with SNR_{eff} of $16\,\mathrm{dB}$, $14\,\mathrm{dB}$ and $11\,\mathrm{dB}$, corresponding to CDF simulation points in Figure 5.5 on the preceding page. Then, Figure 5.6 on the next page shows that the probability of the frequency offset is smaller than 0.05 for SB estimation lies between 0.6 and 0.8. Hence, it becomes clear that higher precision (8 Bits) performs better, with probability larger than 0.9, in all simulated cases.

The channel estimation suffers performance degradation due to the remaining frequency error. In order to quantify this degradation, BER simulations were performed for an LTE uncoded SISO channel under $20\,\mathrm{MHz}$ transmission bandwidth. The simulations are shown in Figure 5.7 on page 64 with normalized frequency error of 0, 0.05 and 0.1.

By observing Figure 5.7 on page 64, it is noticed that the loss due to frequency error is negligible. Moreover, the precision of the SB estimator can be improved by averaging. However, this comes at the cost of increasing latency and consequently energy consumption, as the estimator needs to process multiple symbols in order to produce a single estimate. This is shown in Figure 5.8 on page 64 where the estimator performance for 8-bits using one symbol is compared to SB for various numbers of symbols.

It is also worth noticing that LTE contains specific PSS intended for time synchronization and acquisition of base-station ID. These signals are pseudo-random sequences transmitted in intervals of $0.5\,\mathrm{ms}$, see Section 3.2 on page 29 for more info on LTE specification. SB estimation can also be used to estimate time and frequency in such sequences as it is demonstrated in [48]. However, in the current analysis the CP is used as means of synchronization in order to demonstrate the generalization to any OFDM system.
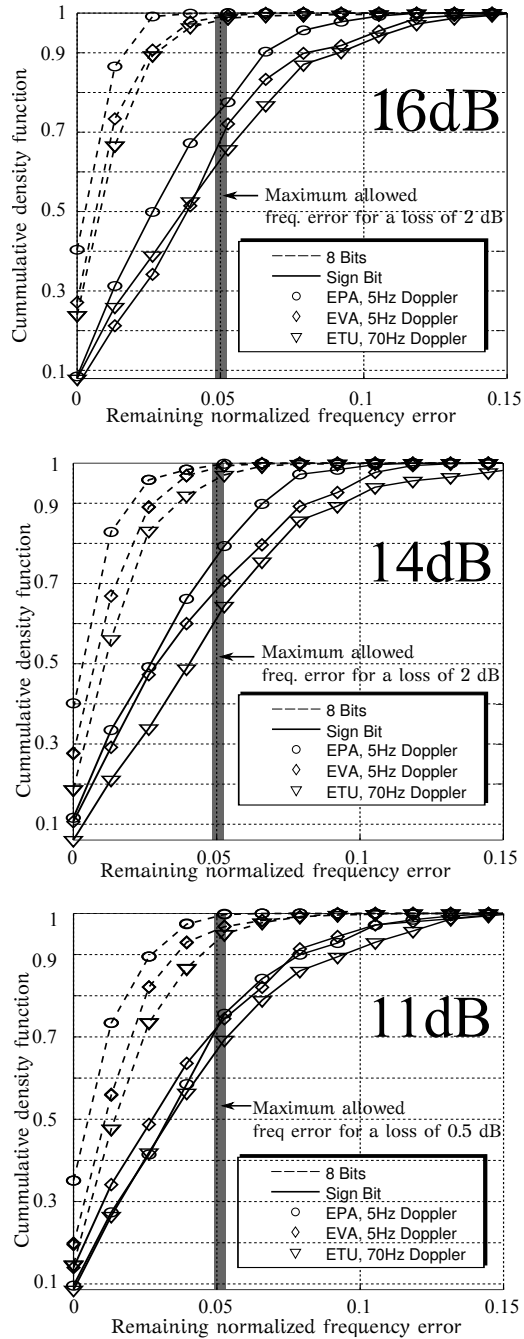
**Figure 5.6.**: Cumulative density function of maximum allowed fractional frequency error normalized to the sub-carrier spacing ($\Delta_f = 15\,\text{kHz}$) for SNR_{eff} of: 16, 14, and 11 dB.
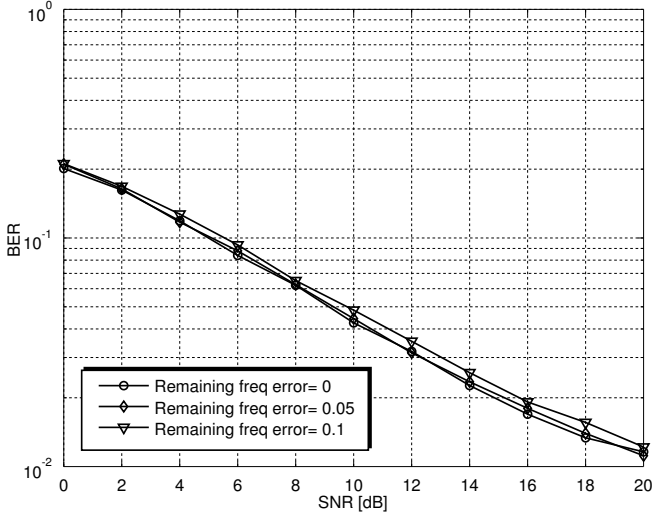
Figure 5.7.: BER for a LTE SISO with 20 MHz bandwidth, QPSK
modulation uncoded transmission with normalized fre-
quency error equivalent to 0, 0.05 and 0.1
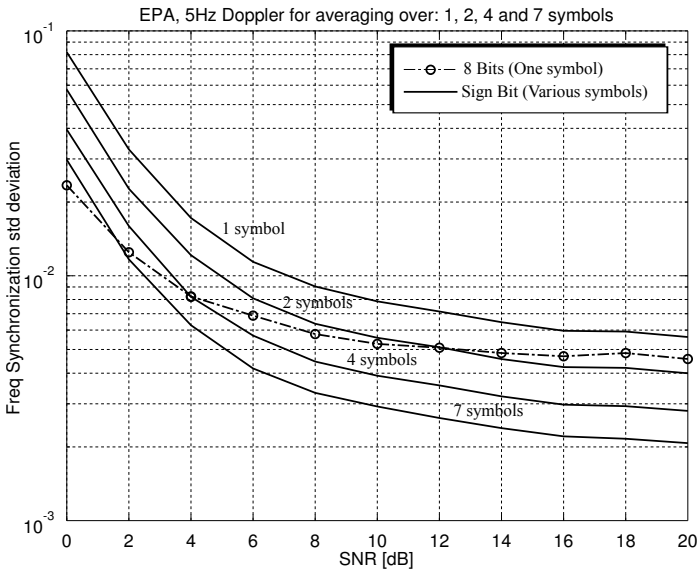
Figure 5.8.: Frequency estimation standard deviation and improved
standard deviation for SB-estimation.

6

Sign-Bit signal to noise ratio

This chapter follows the same procedure as the previous one by first presenting the derivation in full precision format, then introducing the derivation of the SB estimation, and finally providing some performance metrics under LTE scenarios.

6.1. FULL PRECISION ESTIMATION

The CP can be used to estimate the noise in a received signal. This needs the assumption that all communication anomalies have been compensated and that the only source of disturbance is AWGN and thus, i.e., $\tau = 0$, $h(n) = \delta(n)$, and $\Delta\varphi = 0$. As a result, the subtraction of each received sample from its counterpart within an OFDM symbol can be expressed as follows,

$$
\begin{aligned}
\Upsilon(n) &= r(n) - r(n - N) \\
&= s(n) + w(n) - s(n - N) - w(n - N),
\end{aligned}
\tag{6.1}
$$

where $s(n)$ and $w(n)$ denote the signal and noise sample components, respectively. To this effect, there are two cases to be considered: i) $s(n) = s(n-N)$, which holds when $s(n-N)$ is in the CP and $s(n)$ is at the end of the OFDM symbol; ii) $s(n) \neq s(n - N)$, which holds when the two samples belong to different OFDM symbols. This is illustrated in Figure 6.1 on the following page where the first case corresponds to the positions enclosed within the dashed lines with lowest remaining energy. The amount of remaining energy within the dashed-lines is directly proportional to the noise power, since $s(n)$ and $s(n - N)$ in (6.1) cancel each other out, and can be thus used in SNR estimation.

Both the desired signal and the noise terms follow a complex Gaussian distribution. Therefore, $\Upsilon(n)$ also follows a complex Gaussian distribution
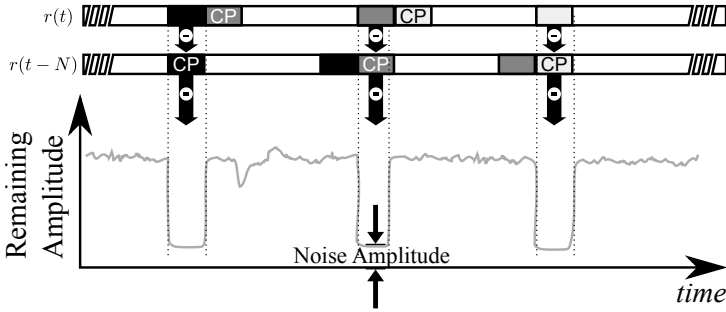
Figure 6.1.: Graphical representation subtraction $|\Upsilon(n)|$ with channel $h(n) = \delta(n)$, $\tau = 0$, and $\Delta\varphi = 0$.

and its amplitude is Rayleigh distributed [49]. Hence, assuming $E[s(n)^2] = \sigma_s^2$, and $E[w(n)^2] = \sigma_w^2$, and taking into account the aforementioned possible cases, the corresponding variance becomes,

$$E[|\Upsilon(n)_{\in CP}|^2] = 2\sigma_w^2, \tag{6.2}$$

and

$$E[|\Upsilon(n)_{\notin CP}|^2] = 2(\sigma_s^2 + \sigma_w^2), \tag{6.3}$$

where the subscripts $_{\in CP}$ and $_{\notin CP}$ indicate whether n is within the CP or not. By combining (6.2) and (6.3), and approximating the expectation as an average over a suitable number of samples, the full precision SNR estimator can be expressed as follows,

$$\widehat{SNR}_{FP} = \frac{\sum |\Upsilon_{\notin CP}|^2}{\sum |\Upsilon_{\in CP}|^2} - 1, \tag{6.4}$$

The full precision SNR estimation technique compares the magnitude of pure-noise samples and noise-plus-signal samples for quantifying the respective SNR.

6.2. SIGN-BIT ESTIMATION

Another application of SB estimation in OFDM systems is low complexity SNR estimation based on the CP. Under the assumption that the only source of disturbance between transmitter and receiver is AWGN, i.e., the same initial scenario as defined for full-precision, the results from subtracting each receiver sample from its counterpart in the CP is expressed as,

$$\Upsilon^q(n) = r^q(n) - r^q(n - N). \tag{6.5}$$
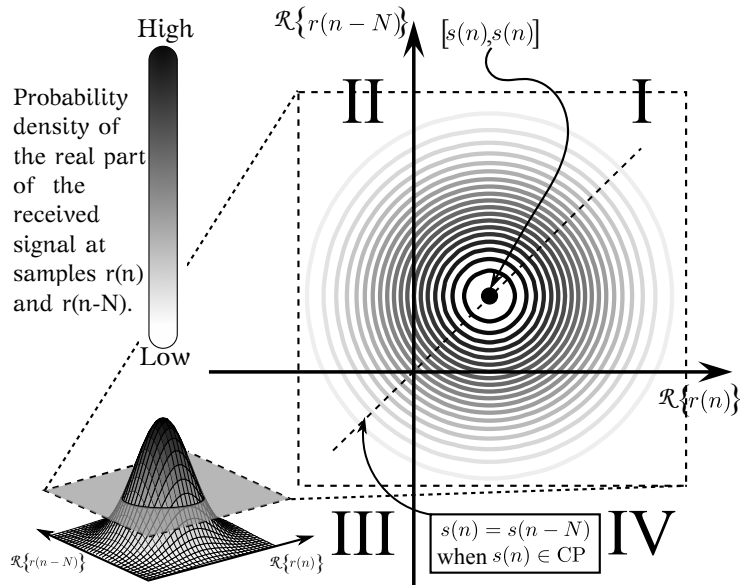
Figure 6.2.: Level curves of the real part of the received signal density at samples $r(n)$ and $r(n-N)$ with the four quadrants indicating the sign of the received sample. The samples $r(n)$ and $r(n-N)$ are in the CP and its counterpart in the symbol, respectively, the same is valid for the imaginary part.

Intuitively, if the noise power is small in comparison to the signal power, the sign of the received sample in the CP is the same as its counterpart within the symbol, resulting in $\Upsilon^q(n) = 0$. However, as the noise power grows, there is an increased risk that the sample in the CP has the opposite sign than in the OFDM symbol and consequently the result of the subtraction yields $\Upsilon^q(n) \neq 0$ probabilistically. This is illustrated in detail in Figure 6.2 which shows curves of the Probability Density Function (PDF) of the real part (the same analysis applies for the imaginary part) of the received signal at samples $r(n)$ and $r(n-N)$, which correspond to the CP and its counterpart in the symbol, respectively. The darker rings indicate higher probability density, while it is also noted that the number of rings and its density are a function of the noise variance. The center of the probability density function is located at the noiseless sample $[s(n), s(n)]$. The dashed line depicts all possible cases where $s(n) = s(n-N)$ with slope equal to unity. Therefore, the probability density of the received sample is always centered on the dashed line.

Next, consider the case in Figure 6.2 where the received noiseless signal is positive and the noise density expands along the four quadrants, e.g., any

ring of the probability density function that occupies all quadrants. It is graphically shown that when subtracting $r^q(n)$ from $r^q(n-N)$, the resulting sign lies within one of the four possible quadrants. The quadrants describe the following four states: i) when no sign change takes effect (I); ii)when only the sign of the $r(n)$ is changed (II); iii) when the both signs are changed (III); iv when only the sign of the $r(n-N)$ is changed (IV). It is noted here that it is impossible to practically distinguish between I and III, since the result of the subtraction yields the same outcome. Likewise, by considering only cases II and IV it is possible to describe the probability density that the subtraction will yield a value different than zero. If we consider that the probability of changing the sign of either sample as

$$P\{sign[s(n)^q] \neq sign[r(n)^q]\} = Q\left(s_n/\sigma_w\right), \tag{6.6}$$

then, it follows that the probability density that the subtraction will yield a value different than zero is given by

$$P\{\Upsilon^q(n) \neq 0]\} = 2Q(s_n/\sigma_w)(1 - (Q(s_n/\sigma_w), \tag{6.7}$$

In order to derive the p (the probability of a subtraction different than zero), all possible values the signal s have to be considered, so that p becomes,

$$p = \frac{4}{\sqrt{2\pi}} \int_0^\infty Q\left(\frac{s}{\sigma_w}\right) e^{-1/2s^2} ds - \frac{2}{\sqrt{2\pi}} \int_0^\infty Q^2\left(\frac{s}{\sigma_w}\right) e^{-1/2s^2} ds. \tag{6.8}$$

The derivation of a closed-form expression[1] for p is subject to analytic evaluation of the integral in the form,

$$\int_0^\infty e^{-ax^2} Q(bx) dx, \tag{6.9}$$

which according to [50] is equivalent to,

$$\int_0^\infty Q(ax) \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} dx = \frac{1}{4} - \frac{1}{2\pi} \tan^{-1}(a) = \frac{1}{4} - \frac{1}{2\pi} \tan^{-1}\left(\sqrt{SNR}\right), \tag{6.10}$$

when $a = 1/\sigma_s$, and $SNR = 1/\sigma_s^2$. Using this result, the first integral in (6.8) can be directly expressed as,

$$4 \int_0^\infty Q\left(\frac{s}{\sigma_w}\right) \frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}} ds = 1 - \frac{2}{\pi} \tan^{-1}\left(\sqrt{SNR}\right) \tag{6.11}$$

---

[1]The author would like to express his gratitude to Dr. Paschalis C. Sofotasios from Tampere University, for the derivation of the closed-form expression of p.

Therefore, it follows that the full equation (6.8) can be rewritten as

$$p = 1 - \frac{2}{\pi}\tan^{-1}\left(\sqrt{\text{SNR}}\right) - 2\int_0^\infty Q^2\left(\frac{s}{\sigma_w}\right)\frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}}ds. \tag{6.12}$$

The derivation of a closed-form expression for p is subject to analytic evaluation of this integral. To this end, the identity of erfc can be used, i.e.,

$$Q(x) \triangleq \frac{1}{2}\text{erfc}\left(\frac{x}{\sqrt{2}}\right). \tag{6.13}$$

As a result, the integral in (6.12) can equivalently re-written as follows,

$$2\int_0^\infty Q^2\left(\frac{s}{\sigma_w}\right)\frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}}ds = \frac{1}{2\sqrt{2\pi}}\int_0^\infty \text{erfc}^2\left(\frac{s}{\sqrt{2}\sigma_w}\right)e^{-\frac{s^2}{2}}ds. \tag{6.14}$$

Furthermore, according to [51],

$$\int_0^\infty e^{-ax^2}\text{erfc}(bx)\text{erfc}(cx)dx = \frac{1}{\sqrt{\pi a}}$$

$$\left[\frac{\pi}{2} - \tan^{-1}\left(\frac{b}{\sqrt{a}}\right) - \tan^{-1}\left(\frac{c}{\sqrt{a}}\right) + \tan^{-1}\left(\frac{bc}{\sqrt{a\left(a + b^2 + c^2\right)}}\right)\right], \quad \tag{6.15}$$

which for $b = c$ reduces to

$$\int_0^\infty \text{erfc}^2(bx)e^{-ax^2}dx =$$

$$\frac{1}{\sqrt{\pi a}}\left[\frac{\pi}{2} - 2\tan^{-1}\left(\frac{b}{\sqrt{a}}\right) + \tan^{-1}\left(\frac{b^2}{\sqrt{a\left(a + 2b^2\right)}}\right)\right]. \tag{6.16}$$

By setting, $a = 1/2$, and $b = 1/\sqrt{2\sigma_w^2}$, it follows that

$$\frac{1}{2\sqrt{2\pi}}\int_0^\infty \text{erfc}^2\left(\frac{s}{\sqrt{2}\sigma_w}\right)e^{-\frac{s^2}{2}}ds =$$

$$\frac{1}{2\pi}\left[\frac{\pi}{2} - 2\tan^{-1}\left(\frac{1}{\sigma_w}\right) + \tan^{-1}\left(\frac{1}{\sqrt{\sigma_w^4 + 2\sigma_w^2}}\right)\right]. \tag{6.17}$$

By noting that $\text{SNR} = 1/\sigma_w^2$ and carrying some basic algebraic manipulations one obtains,

$$2\int_0^\infty Q^2\left(\frac{s}{\sigma_w}\right)\frac{e^{-\frac{s^2}{2}}}{\sqrt{2\pi}}ds =$$

$$\frac{1}{2\pi}\left[\frac{\pi}{2} - 2\tan^{-1}\left(\sqrt{\text{SNR}}\right) + \tan^{-1}\left(\frac{\text{SNR}}{\sqrt{1+2\text{SNR}}}\right)\right]. \qquad (6.18)$$

Finally, by substituting (6.18) into (6.8), the following closed-form expression for the p is deduced,

$$p = 1 - \frac{2}{\pi}\tan^{-1}\left(\sqrt{\text{SNR}}\right) -$$

$$\frac{1}{2\pi}\left[\frac{\pi}{2} - 2\tan^{-1}\left(\sqrt{\text{SNR}}\right) + \tan^{-1}\left(\frac{\text{SNR}}{\sqrt{1+2\text{SNR}}}\right)\right], \qquad (6.19)$$

which after some manipulation reduces to

$$p = \frac{3}{4} - \frac{1}{\pi}\tan^{-1}\left(\sqrt{\text{SNR}}\right) - \frac{1}{2\pi}\tan^{-1}\left(\frac{\text{SNR}}{\sqrt{1+2\text{SNR}}}\right). \qquad (6.20)$$

Then, the SNR and p are linked by,

$$\widehat{\text{SNR}}_{\text{SB}} = 1 + \frac{\sqrt{1+\tan^2(\pi p)}}{\tan^2(\pi p)}$$

$$= 1 + \frac{\cot(\pi p)}{\sin(\pi p)}. \qquad (6.21)$$

The value of p is approximated by averaging enough number of samples $|\Upsilon^q_{\in\text{CP}}|$ so that a reliable SNR estimate can be computed. Notably, even though the presented analysis considers only the AWGN scenario, it is subsequently shown that SB estimation goes beyond this condition. Moreover, the advantages offered by these techniques when decreasing the corresponding hardware complexity are presented extensively.

6.3. ESTIMATION PERFORMANCE

A mobile terminal pursues the best possible connectivity by continuously seeking a base-station with better transmission conditions. Since the mobile terminal is already connected to the network while searching for other base-stations, some properties of the environment are unchanged during the transition, since the terrain cannot change instantly. Therefore, it is possible to assume some knowledge of the currently most suitable channel model.

The average estimated SNR is depicted in Figure 6.4 on page 73 where a bias is observed in both methods, distance between the mean and the ideal value, but is more significant in the 8-bit estimation. The bias results from the channel's impulse response affecting samples in the CP. Hence, the larger the

channel delay spread, the more samples are affected and thus, the larger the corresponding bias. This can be seen in the figure by looking into EPA (the shortest delay spread), and ETU (longest delay spread), where the estimated value is the closest to the ideal in EPA and most distant from it in ETU channel model.

Samples distorted by the channel impulse response affect the performance of the SB only by its sign, not by its amplitude, contrary to the 8-bit method. Thus, it can be graphically deduced that for any channel model, the bias in the SB method is smaller than the 8-bit. Furthermore, it is possible to combat the channel dependent bias by decreasing the size of the moving sum and consequently reducing the number of samples introduced to the estimation window that were affected by the channel impulse response. However, this in turn increases the estimator's variance, which affects its performance.

Another way to remove the bias is to characterize it statistically for each method and each channel model and then compensate the estimation accordingly. This can be possible under the assumption that some knowledge of the channel properties exists. The results of this compensation are depicted in Figure 6.3 on the following page with the mean and 5%-95% confidence intervals of the estimation methods. The region with solid lines corresponds to SB estimation while the region with the dashed lines corresponds to 8-bits estimation for EPA, EVA and ETU. The centered line represents the mean value after compensation and the envelope around it denotes its confidence interval. From the figure, it can be seen that for all channel models the 8-bit has the smallest variance, ie approximately $\pm 2\,$dB. The variance of SB is slightly larger than the 8-bit one in most cases with the worst case scenario being EPA, i.e., with a variance of $\pm 4\,$dB at true SNR of $16\,$dB. The figure depicts the result of simulations performed with 21 OFDM symbols to create a single estimate, the variance -or confidence interval- can be reduced for this worst case by increasing the number of symbol used per estimate, at the cost of increasing estimation latency. It is noted here that the compensation for SB has much lower complexity in terms of hardware compared to the 8-bit. This is achieved because the SB-SNR compensation requires only one Look-Up Table (LUT) per channel model, whereas 8-bit-SNR compensation needs one LUT per channel model with higher resolution, in addition to a division module, which is typically a costly operation in hardware systems.
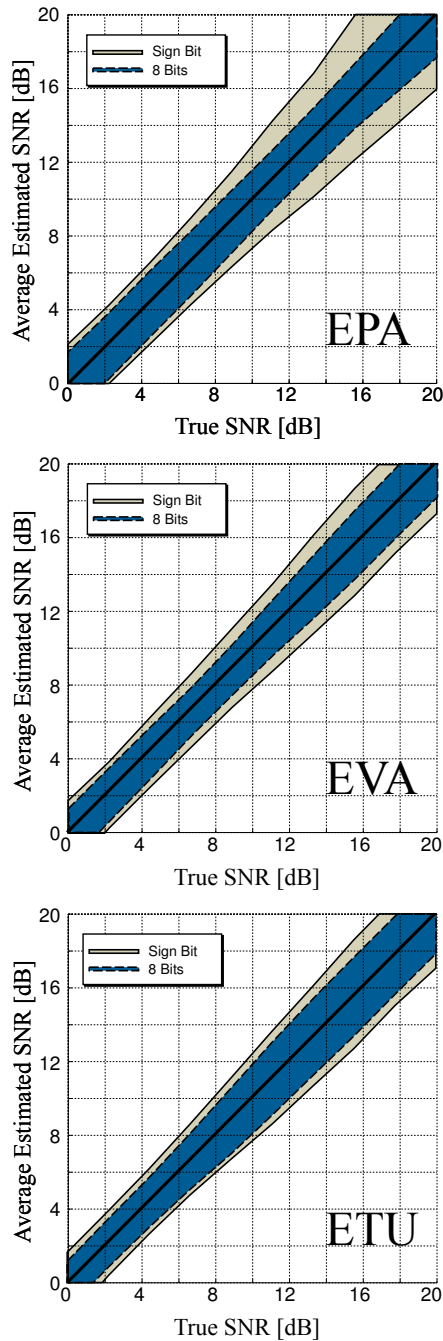
Figure 6.3.: Estimated SNR with 5% - 95% confidence intervals, compensated for channels EPA, EVA and ETU. Each estimation is realized as an average of 21 consecutive OFDM symbols with a moving sum size of 115 samples.
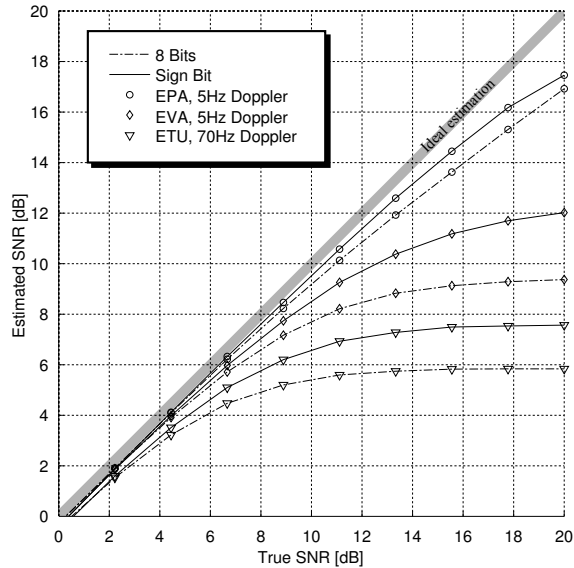
Figure 6.4.: Average estimated SNR uncompensated simulated for EPA, EVA, ETU, and AWGN channel models. Each estimation is realized as an average of 21 consecutive OFDM symbols with a moving sum size of 115 samples.

So far, the received signal used for SNR estimation is assumed to have no CFO, which requires perfect CFO compensation. In this case the remaining frequency error affects the SNR estimation. This is illustrated in Figure 6.5 on the next page, where it can be seen that the estimated SNR deviates from the true SNR when there is any CFO remaining in the signal. As an example, when the remaining normalized frequency error is 0.05, according to Figure 5.6 on page 63, the probability of the error being smaller than 0.05 for an SNR_{eff}=11 dB is larger than 70% for all cases. This implies that even in the presence of a remaining frequency error in the SNR estimation, this rarely reaches a value that can severely affect SNR estimation.

In this section the performance evaluation of SB estimation techniques have been presented in the context of an LTE receiver with realistic channel models. From the simulations, it can be concluded that the SB estimation has similar performance to 8-bit, for symbol timing and SNR. Notably, the CFO estimation, even though the SB estimation performance is worse than 8-bit, it rarely appears to be harmful enough for the subsequent tasks in the receiver.
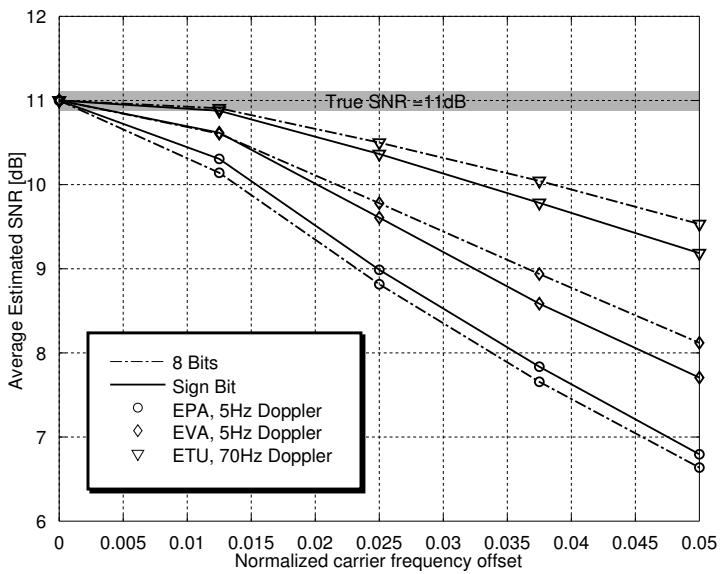
Figure 6.5.: Effect of remaining CFO into SNR estimation for EPA; the CFO is normalized to the sub-carrier spacing (15K Hz).

7

Complexity analysis and proof of concept

The complexity analysis is performed in terms of area and power of the hardware architecture. In order to ensure consistency with performance evaluations, an equivalent architecture with an 8-Bit resolution is created and compared to the SB architecture. This is illustrated in Figure 7.1, where white blocks are considered external and irrelevant to the complexity analysis. The dashed line in figure, and in subsequent architectural diagrams, represents the path where a single real value is passed. On the contrary, the solid lines represent I and Q components i.e., two values, passing between the blocks. A more detailed description of each of the architectures is additionally depicted in Figure 7.2 on the following page. Importantly, both architectures can estimate the peak-based CFO using the external arctan module which is typically implemented with a COordinate Rotation DIgital Computer (CORDIC) [52] [53] in Figure 7.1. However, only the SB-architecture appears to be capable of estimating the angle-based and peak-based CFO. This is due to the much higher complexity requirements that the angle-based CFO imposes when the signal is quantized to 8-bits with no performance benefits, in comparison with the already available 8-bit peak-based estimation.
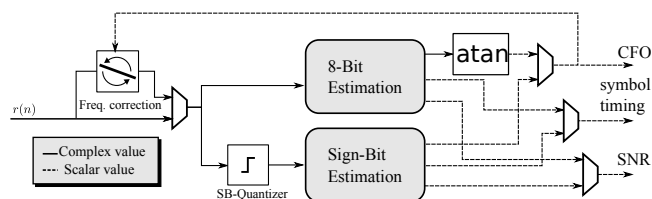
Figure 7.1.: Complexity analysis conceptual image, where frequency correction, SB-quantizer, arctangent, and multiplexers, are considered external and outside the complexity analysis.

7.1. EQUIVALENT ARCHITECTURE

The architecture depicted in Figure 7.2a is capable of estimating the intended parameters based on the methodology introduced with infinite precision, but this time its input's signal is quantized to 8 bits. Its operation is based in two modes: firstly, the *time-frequency* mode is in charge of performing initial time and prepares the signal for CFO estimation; and secondly, the *SNR* mode is in charge of performing SNR estimation. Since the *SNR* mode requires input from the time-frequency, the modes are mutually exclusive and can share some hardware components.

7.1.1. TIME-FREQUENCY MODE

In this mode the correlation peak is calculated by multiplying the input signal with its conjugate version delayed by N samples and subsequently adding the previous A multiplications - see equations (5.5) and (5.6). This is realized in hardware by the combination of a First Input First Output (FIFO) memory and a complex multiplier, that calculate the value of $\beta(n)$ at each sample. Subsequently, two moving-sums add the last A multiplication results for I and Q, respectively.
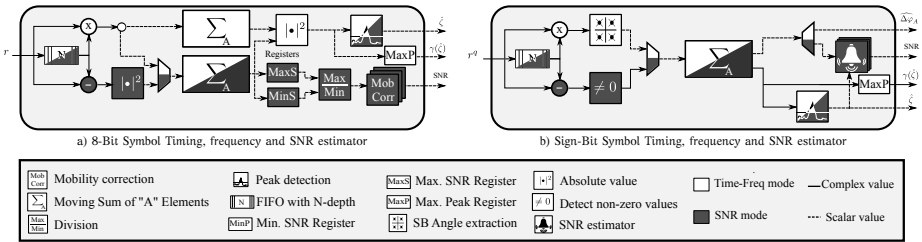
Figure 7.2.: Block diagram of architecture of a) 8-Bit and b) SB hardware architecture, capable of estimating symbol timing, CFO and SNR. The dark modules are used for SNR estimation, the white for time-frequency and the mix-colored for both.

The symbol timing according to equation (5.4) is estimated as the position (in samples) where the correlation is maximum. As a result, a Finite State Machine (FSM) is used for keeping track of the correlation output. It is recalled here that the correlation value is a complex quantity and as a consequence, its absolute value needs to be computed in order to test for maximum amplitude.

7.1.2. SNR MODE

In this mode the SNR is estimated by approximating the amount of energy within the CP and outside the CP. This is accomplished in hardware by

subtracting the input signal from its delayed version. For the delay, the same FIFO memory as in the *time-frequency* mode can be used, whereas the subtraction and the posterior energy calculation is done by a 8-bit subtraction and an absolute-value module that squares the result of the subtraction and averages both I and Q components.

One of the two available moving sums, the bottom one from Figure 7.2 on the preceding pagea in the architecture, can be used for adding the amount of energy within the CP length. The FSM also ensures that the amount of energy corresponding to the energy inside and energy outside the CP is stored in the registers MinS and MaxS, respectively. The SNR is then realized by computing the ratio between MaxS and MinS. The ratio is implemented as a sequential divider that converges in several clock cycles. A correction is then applied to the calculation to account for multi-path channels, as previously discussed in Section 6.3.

7.2. SIGN-BIT ARCHITECTURE

Analogous to the 8-bit architecture, the SB architecture, depicted in Figure 7.2 on the facing pageb can estimate the same parameters while it can additionally provide the angle-based CFO. Furthermore, it operates in the same two modes; first time-frequency and SNR modes, executed in that order. It can also be noticed that the SB architecture has considerably fewer elements than its higher precision counterpart. Additionally, the majority of the elements can be fully shared between the two operational modes.

7.2.1. TIME-FREQUENCY MODE

The heavy quantization plays an important role in reducing the corresponding hardware complexity. The size of the FIFO is reduced dramatically while the multiplication becomes a single xor-boolean operation. In addition, a complex CORDIC architecture can be avoided when extracting the phase of β in (5.16), since the phase can only take four possible angles, namely, $0, \pi/2, \pi$, or $-\pi/2$. The moving-sum is also simplified since it is not necessary to keep two separate moving-sums for I and Q components as these components can be straightforwardly extracted from the angle itself.

Similar to the 8-bit architecture, the peak detection module is an FSM that monitors the value of $|\gamma^q(n)|$ for the previous 2 symbols and determines the location of the corresponding maximum. The FSM then reports this position as the symbol timing estimation $\hat{\zeta}$.

The architecture for time-frequency only was presented in [54] and was later fabricated in a CMOS 65 nm process with the layout illustrated in Figure 7.3 on the next page. The measurements from the fabricated SB-time-frequency-correlator are shown in Figure 7.4 on page 79, where the red and blue color correspond to the simulated and measured CFOs respectively
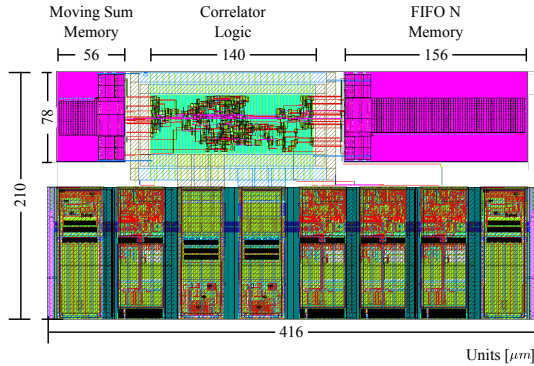
Figure 7.3.: Fabricated SB-time-freq-correlator in ST 65 nm Technology with High VT low power cells.

$(\widehat{\Delta\varphi_a^q})$. Likewise, the cyan and magenta curves correspond to the simulated and measured correlation peak, respectively. They both show a distinctive peak corresponding to the position where the symbol-timing $(\hat{\zeta})$ is located. It is noted that the slight difference between simulation results and measurements results occurs for two reasons: firstly, an initialization phase when the FIFOs are being filled most noticeable in the CFO); secondly, a sample offset between the simulation and the measurements. It is also shown that the measurements report an average power dissipation of 0.62 mW for LTE, with 20 MHz bandwidth, with a power supply applied to the core of 1.2 V.

7.2.2. SNR MODE

In this mode, the incoming signal is assumed to be CFO-clean, i.e., symbol-timing and CFO correction have been performed in the previous mode. The sign of the the input signal is subtracted from its delayed version with the aid of the same FIFO used in *time-frequency* mode for correlation. The block named non-zero detector counts the number of received samples that had changed sign[1] and inputs this value to the moving-sum. Notably, the same moving-sum used for the *time-frequency* mode. The SNR is then estimated by inserting the total number of flipped signs to the estimation tile, as depicted in Figure 7.2 on page 76b with a bell icon, and implemented as a LUT. The LUT extracts the SNR based on the corresponding probability p. Since p varies in relation to the current channel, the architecture contains an estimation tile LUT per channel model.

---

[1]A change of sign is perceived as a non-zero result from the subtraction (See Sect. 6.2)
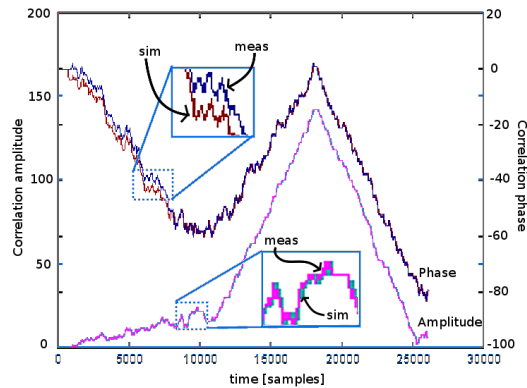
Figure 7.4.: Measured Peak and Angle compared to Matlab simulations. The CFO estimation red and blue correspond to the simulated and measured, respectively. While $|\gamma^q(n)|$ is represented with cyan and magenta for simulated and measured, respectively. The horizontal axis represent number of samples and the vertical axes a integer number relative to the actual value of the corresponding parameter.

7.3. COMPLEXITY TRADE-OFFS

In order to obtain a clear idea of the complexity trade-offs associated with the SB estimation, the SB architecture and its equivalent higher precision counterpart are compared to each-other along with relevant existing techniques for estimating symbol-timing and CFO. The top-level architecture previously described is synthesized, placed, routed and simulated to extract area and power numbers, the top-level layout is shown in Figure 7.5 on the following page. It is noted here that the estimation principle can be applied to any OFDM system, which is demonstrated by designing a rather scalable combined architecture so that it can be conveniently employed in three OFDM standards, namely, WLAN, DVB-H in 4K mode, and LTE. The synthesis results for both SB and 8-bit are depicted in Figure 7.6 on page 81 with the total area and power consumption split into memory and logic contribution.

It can be seen that the total area of the 8-bit architecture is 4 times larger than it's SB equivalent and dissipates 6 times more power. Even though one could expect the 8-bit area overhead to be at least 8 times larger, this is not the case due to the fact that larger memories, in terms of storage, are typically denser. For example, the memory in the 8-bit architecture used for correlation (the largest in the design) has an area density of $0.9\mu m^2$-per-bit for the 8-bit architecture, which is almost a factor of 2 denser than the SB
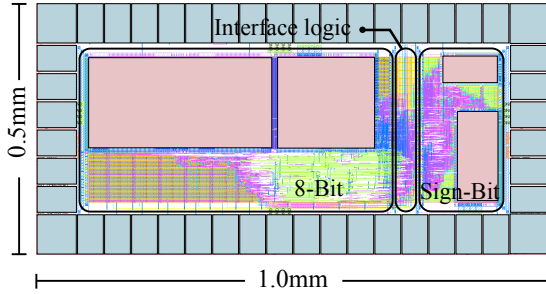
Figure 7.5.: Fabricated SB-time-freq-correlator in ST 65 nm Technology with High VT low power cells.

memory with $1.6\mu m^2$-per-bit.

In the SB architecture, the logic, which occupies 45% of the total area, dissipates 40% of the total power which implies that the power and area are memory dominant. Likewise, in the 8-bit architecture, the logic occupies 29% and dissipates 46% of the total energy. The simulated leakage was excluded from the figure since, according to simulations, it only accounts for 0.04% and 0.03% of the total dissipated power for SB and 8-bit, respectively.

In order to see what other logic modules can be meaningful contributors, a further partition into individual hardware blocks is shown in Table 7.1. Recall that the SNR is calculated differently in both methods, i.e., in the case of the 8-bit architecture, a block to perform sequential division is used (See Section 6.1), and in the case of SB a LUT is used to extract the SNR estimation based on the statistics of the signal subtraction (See Sections 6.2 and 6.3). In both architectures the largest area/power contributors are the modules containing memories. The third most important contributor is a register bank used to store intermediate and final values. This accounts for 14% of area with 6% of the power for the SB and 6% of area with 2% of the power for the 8-bit architecture. It is also noted that the contribution of the correlation FIFO and the moving sum FIFO adds up to a share larger than the one reported in Figure 7.6 on the facing page. In order to avoid using area expensive dual-port memories to perform FIFO functionality, a much more compact single-port, surrounded by extra logic needed to emulate a FIFO, is alternatively used. Furthermore, the moving sum is a combination of a FIFO, arithmetic elements and registers. Thus, each FIFO implementation accounts for more than just memory elements.

Table 7.2 shows the comparison of the proposed architecture to other designs found in literature. It is worth mentioning that [55], [56] and [57] are designed for time-frequency synchronization of WLAN where there is a synchronization preamble. In the WLAN legacy version, the preamble is a predefined sequence consisting of 10 short symbols of 16 samples each, with

		Sign-Bit		8-Bit	
		Area $[mm^2]$	Power $[mW]$	Area $[mm^2]$	Power $[mW]$
■	Memory	0.018	0.288	0.094	1.543
□	Logic	0.015	0.191	0.038	1.314
	Total	0.035	0.479	0.133	2.857
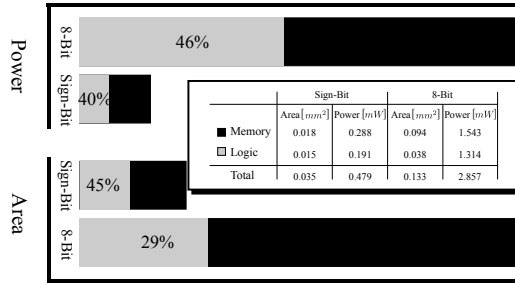
Figure 7.6.: Power and area distributions for SB and 8-bit after place and route. Simulation performed with LTE 20MHz bandwidth and fully-back annotated simulation at 1.2V supply

their values being specifically designed to provide a high correlation peak. The proposed architecture can be used for WLAN, but it was intentionally designed to support more memory demanding standards such as LTE and DVB-H. Thus, a downsized version of the architecture is used for fair comparison, i.e., the FIFO is reduced from 4096 words to 16 and the memory in the moving sum is reduced from 512 to 160 words. For LTE and DVB-H the size of the largest FIFO is the size of its corresponding FFT, which is 2048 and 4096 respectively, whereas in WLAN, the same FIFO is reduced to the size of a short training symbol in the preamble with size 16. It is also seen in Table 7.2 that both architectures have a smaller gate-count and power consumption than previously published works, regardless of the extra-functionality. Furthermore, the SB architecture can reduce the area and power to a tiny fraction of those reported in literature. Besides being more power and area efficient, both architectures provide a coarse SNR estimation that can improve energy efficiency even further, by providing an early channel prediction in later stages.

Table 7.1.: Area and Power breakdown of the various parts of the architecture.

	8-Bit [%]		Sign-Bit [%]	
	Area	Power	Area	Power
Correlation FIFO	49	32	45	42
Moving Sum FIFO	34	45	28	37
Register Bank	6	2	14	6
Peak detection	2	0.67	6	4
Complex multiplier	3	9.7	-	-
SB Complex mult + angle extraction	-	-	0.5	1
SNR Est. (divider + comp.)	4.06	1.101	-	-
SNR Est. (LUT)	-	-	2	0.01
Remaining logic	2.3	9	4	10.5

Table 7.2.: Comparison to other implementations in literature, where all technology processes were scaled to 65 nm.

Design	8-Bit*	Sign-Bit*	[55]	[56]	[57]
Technology [nm]	65	65	180	250	180
Quant. bits	8	1	10	10	10
Gate count	30K	8K	37K	-	104K
Vdd Core [V]	1.2	1.2	-	2.5	-
Area [mm^2]	0.06	0.02	-	0.42	0.196
Power [mW]	1.4	0.2	14	13.5	26.5

Scaling rules: A~1/s^2 and P~(1/s)(Vdd/Vdd'). Vdd for 180 nm assumed to be 1.6 V.

*Down-scaled architecture to support only WLAN for fair comparison. In addition capable of estimating SNR.

8

Summary of Part I

In the current part it has been demonstrated that it is possible to estimate symbol timing, Carrier Frequency Offset (CFO), and Signal-to-Noise Ratio (SNR) only with the use of the received Sign-Bit (SB). It has also shown its advantages and disadvantages in terms of hardware complexity and performance. Performance simulations under LTE typical scenarios indicate little or no degradation for symbol-timing and CFO estimation in comparison to an equivalent 8-bit architecture.

Based on simulations, it was found that for SNR estimation the performance of SB is similar to its higher precision equivalent with much lower complexity. The potential of SB estimation techniques for low power receiver terminals was demonstrated with post-layout power simulations with back-annotated information. The simulations showed that the SB-estimation architecture dissipated at least 1/70 of the power consumption compared to similar existing works in the open technical literature.

Additional measurements of the time-frequency correlator have been presented and showed an average power consumption 0.62 mW for LTE with 20 MHz bandwidth and power supply of 1.2 V. Finally, the applicability to more than one OFDM standard is shown by including enough hardware resources in the fabricated architecture to be able to provide service to LTE, WiFi, and DVB-H.

Part II.

Custom DSP processing for LTE-A

This part comprises the work carried out during a 6-month exchange to Inter-university Microelectronics Research Center (IMEC) in Leuven, Belgium. The author was invited to carry a feasibility study regarding band extraction, also referred as channelization, in LTE-A with Carrier Aggregation (CA). The outcome of this study is published in the following international conference:

- Diaz, I., Torrea-Duran, R., Pollin, S., Van der Perre, L., and Öwall, V., "Selective Channelization on an SDR Platform for LTE-A Carrier Aggregation," *IEEE International Conference on Electronics, Circuits, and Systems*, October, 2012.

9

Common filtering methods and target architecture

The increasing number of mobile applications and services had urged the ITU, more specifically, the radio-communication sector, to call for the development of a new technology that can cope with future user demands in terms of security, quality of service, and transmission data-rate.

The 3GPP has proposed LTE as the fourth generation of mobile communication technologies. LTE-A addresses and even surpasses the IMT-Advanced requirements set by ITU-R, more details about LTE specifications are given in Chapter 3. LTE-A makes use of state-of-the-art technology components, such as wider-band transmission up to 100 MHz with Carrier Component Aggregation (CCA), multiple-antennas with coordinated multi-point, and denser infrastructure of relaying solutions [58]. Among the previous mentioned, CCA represents one of the major challenges for the design of mobile terminals.

On the hardware side, Software Defined Radio (SDR) offers a compelling alternative as the access platform for next generation mobile terminals, where energy efficiency is combined with a high degree of reconfigurability and adaptability [59]. One incentive for the development of SDR based platforms is the fact that there is an increasing number of wireless standards in mobile applications, i.e., WiFi, DVB-H, WiMAX, Bluetooth, LTE, *etc*. SDR platforms offer the possibility to support any current and possibly any future wireless technology by software upgrades, thus enhancing the productivity life-span of a mobile platform and consequently decrease the development costs [60].

The SDR platform used for this study is formally named Architecture for Dynamically Reconfigurable Embedded Systems (ADRES), and is an architecture template that targets power efficient and high dynamic applications by combining features from Very Long Instruction Word (VLIW) and Coarse Grain Reconfigurable Architecture (CGRA) [61].

In this thesis the feasibility of reusing the reconfigurability of ADRES
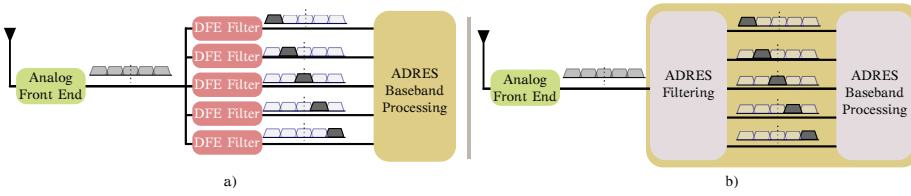
Figure 9.1.: Conceptual view on the two alternatives to combine LTE-A channelization with a SDR architecture.

architecture for means of channel band selection is explored. The most demanding aggregation scenarios by LTE-A, i.e., scenarios 2, and 3 in Figure 3.6 on page 33, require to filter or channelize 5 20MHz channels concurrently, this can be done by building 5 completely independent DFE filters adjusted to reject all but one selected 20 MHz channel each, however, since the aggregation scenario is not known a priori, designing for all possible scenarios might result in an overwhelming overhead. If instead the filtering is done by a programmable SDR architecture, this overhead might be reduced. These two alternatives are illustrated in Figure 9.1, it is important to note that the ADRES platform used is not duplicated, as it might seem from the figure, but rather, the functionality of the ADRES is extended to include ADRES filtering, in addition to other baseband operations.

In this chapter a basic concepts of some common digital filtering methods are introduced. This serves as basis to understand the costs imposed into the ADRES architecture. Furthermore, a brief description of the different parts of ADRES is also presented in this chapter.

9.1. COMMON DIGITAL FILTERING METHODS

Digital filters have a large number of applications, there are many books written about its design and implementation, it is an active research area and new methods are being published all the time, thus, the author's intention is not describing in details all the choices available, but rather introduce the few methods used in for this thesis, i.e., Finite Impulse Response (FIR), Cascaded Integrator-Comb (CIC) filters, commonly used in multiple-rate applications, and Fast Fourier Transform (FFT) as an alternative way to perform the filtering.

9.1.1. FIR FILTERING

Filters are typically classified in terms of its impulse response, i.e., Infinite Impulse Response (IIR) containing an active feedback path, and Finite Impulse Response (FIR) with no feedback associated. As its name implies in FIRs, a single output sample depends on a finite number of input's samples

given by the filter's order, whereas in the case of IIR, it depends on an infinite number of input samples. An FIR can be visualized as a convolution sum, even though they can also be implemented as a special Discrete Fourier Transform (DFT) case [62]. Among the advantages of FIR filters one can list the following:

- they can be designed to have a linear phase by making the coefficient sequence symmetric,

- FIR filters can cope well with quantization, as they have no feedback, rounding errors can be avoided if desired,

- and they are always stable, due to the lack of feedback.

The output samples of an FIR are given by,

$$y(n) = \sum_{k=0}^{N} c_k x(n-k),$$ (9.1)

with its corresponding transfer function expressed as

$$G(z) = \sum_{n=0}^{N} c_n z^{-n}.$$ (9.2)

where it can be seen that the output will always be bounded by the values of the coefficient sequence c, and previous N input values, Figure 9.2 illustrates a typical unfolded architecture of an FIR filter, where the boxes labeled as D represent a single sample delay and the c_i is the ith sample of the impulse response sequence, commonly referred to as filter coefficients. Quantization errors in fixed-point implementation can be completely avoided by accounting for worst case scenario during the computation of the filter's output.
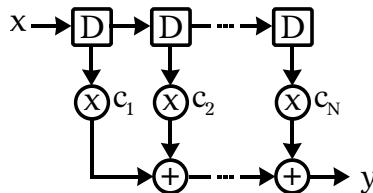
Figure 9.2.: Unfolded FIR filter in direct form.

The unfolded version of this direct-form FIR can be improved with respect to silicon area by folding, and with respect to speed by pipelining. The same filter can be implemented in a transposed-form, which also increases speed, at a cost of increases silicon area.

9.1.2. CIC FILTERING

CIC filters are a special case of low-pass filters consisting of two main stages, i.e., integration stage and comb stage. The order of the stages changes the filter behavior so that it can be used as decimator (when integrator stage is followed by the comb stage), or as interpolator (when the comb stage is followed by the integrator stage), an intermediate process of removing samples (down-sampling), or adding zeros (up-sampling), can be placed in between the two stages to decrease/increase sampling rate according to the intended application.
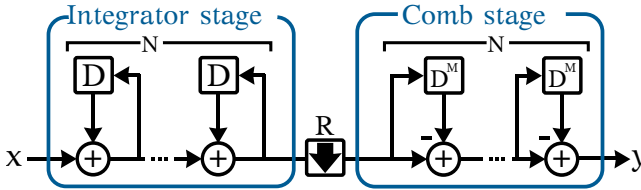
Figure 9.3.: Unfolded decimation filter conceptual diagram.

Figure 9.3 depicts the general structure of a CIC decimation filter, i.e., the integrator stage is followed by a comb stage with the down-sampling module depicted with an arrow pointing down and an R denoting the decimation factor. As it can be seen from the figure, each integrator represents a single-pole IIR filter with a unity feedback coefficient, whereas each comb module is an odd-symmetric FIR filter with M delays at each comb module. M is called differential delay and can be any positive number, but it is usually restricted to 1 or 2 [63]. Here N is the total number of modules in both stages.

The frequency response of the decimation filter is given by,

$$G(z) = H_{int}^N(z)H_{comb}^N(z) = \frac{(1 - z^{-RM})^N}{(1 - z^{-1})^N} = \left[\sum_{k=0}^{RM-1} z^{-k} \right]^N, \qquad (9.3)$$

which shows that even though the CIC filter contains IIR elements, the total frequency response is equivalent to that of N FIR filters, with all the characteristics of FIRs, it is also important to note that an CIC filter always behaves as a low pass filter whose frequency can be adjusted by modifying N, M, and R.

The main disadvantage of CIC filters is a poor filter performance. The frequency response of a typical CIC filter has non-flat pass band and a very long transition band. Hence, compensation filters are often used in combination with CIC filters to improve frequency response.
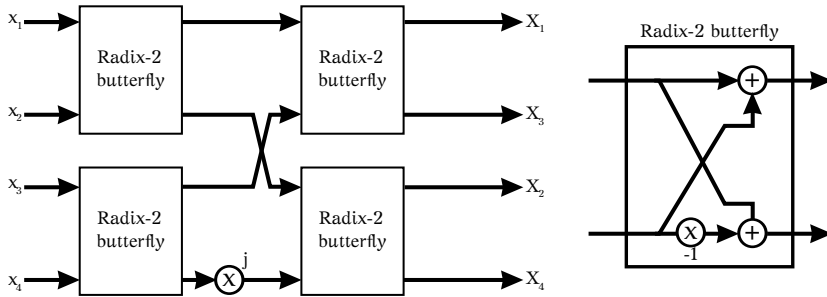
Figure 9.4.: 4-point FFT fully unfolded architecture, with reversed output (decimation in frequency).

9.1.3. FFT-BASED FILTERING

The FFT, as the names implies, transfers a signal representation from time to frequency and vice versa. An FFT with the appropriate resolution can be used an alternative to filtering, since all frequency components are present after the signal's transformation. The popularity of FFT is due to the complexity reduction that represents when compared to a regular DFT implementation, i.e., from $\mathcal{O}(N^2)$ to $\mathcal{O}(N\log_2(N))$. The complexity reduction can be visualized from the fact that the FFT can use a divide-and-conquer approach. It computes intermediate values that can be re-used in later stages, and thus, the total number of operations is reduced. The definition of the DFT is given by,

$$X(n) = \sum_{k=0}^{N-1} x(k)e^{2\pi nk/N}, \quad n = 0, \cdots, N-1. \tag{9.4}$$

The complexity reduction is achieved by assuming N is a highly composite number and a factor of two i.e., $N = 2^m$, where m is a positive integer [64]. A large FFT can be implemented relying on simple components such as adders and multipliers, the most important element of an FFT architecture is what is commonly known as a butterfly unit, it represents the most common and simplest element in an FFT module with any size. A simple fully parallel 4-point FFT and a butterfly unit are depicted in Figure 9.4, note that the multiplications performed are, a simple sign change and multiplication by j, the latter in practice is just a logical operation and can be done without any additional circuitry.

The figure shows a fully unfolded FFT, even though this architecture is the fastest, it is not always a practical or even desired implementation, instead the architecture can be folded, or time-multiplexed or pipelined, where area is traded-off with speed [65].

9.2. ADRES ARCHITECTURE

Digital circuitry drives the potential for highly flexible circuits, capable of being controlled by software instructions. Today there is a clear distinction, in digital design, between application specific hardware and application generic hardware. The first is designed with a unique purpose, but performs poorly in any other context. The latter is designed for a large variety of applications. Even though generic hardware is not as efficient as a specific hardware, for any single application, it provides many advantages in terms of re-usability, flexibility, and costs. SDR takes advantage of these properties and focus them in the wireless communication field.

The goal of an SDR platform is to replace all analog components in a radio transceiver with digital signal processing equivalents. While, current technology has not been able to reach this goal, the number of analog elements in the terminal has been dramatically reduced. An SDR platform achieves high flexibility and high performance by utilizing a number of interconnected circuits such as, analog and digital front-ends, specialized digital accelerators, a digital processor, *etc*.

In this study the ADRES SDR architecture is used for simplification of the digital front-end when LTE-A is the target standard. The filtering required for LTE-A channelization is moved from a fixed digital front-end to the more flexible part of the SDR receiver.

9.2.1. ADRES TEMPLATE

ADRES is an architecture template that can be easily configured for a large variety of applications. For example in [66], an ADRES template serves as transceiver for various OFDM standards by distributing the required processing tasks across a heterogeneous Multi-Processor System on Chip (MPSOC). In this MPSOCs implementation, sub-functions are divided according to their nature (control or digital signal processing). The heterogeneous MPSOC contains a digital baseband FE for packet detection. It also contains two baseband processors for data processing. And the final element is an ARM926 used for event-based control flow, inter-core data transfer, and memory access control. This combination of different architectures in a single SoC is key for a balanced platform in terms of performance and energy consumption.

Since the basic template can be optimized for any specific application, it is reasonable to assume that the optimal CGA implementation for channelization of each of the CA scenarios result in different ADRES instances with specific functionality specialized for each of the aggregation scenarios. However, even if these instances are optimal for each aggregation scheme, they might not be optimal when considering also the remaining elements of the baseband. Therefore, instead of building various specialized instances, it was decided to map the channelization functionality into one instance. This instance is optimized for the entire baseband operation rather than just filter-
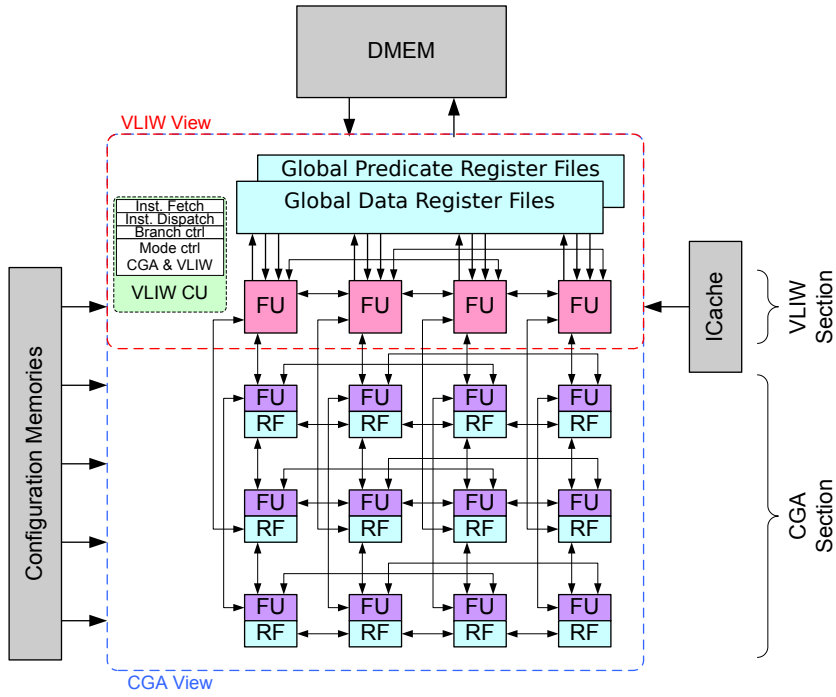
Figure 9.5.: ADRES instance example with 4x4 functional units, more details of this architecture can be found in [61].

ing. Then, the overhead imposed by the additional functionality is analyzed to find the best filtering alternative.

One more example of how ADRES can be adapted to many applications can be seen in [67], where an ADRES instance is used for supporting standards such as IEEE 802.11g/n with 64QAM in 20MHz bandwidth and LTE with 16QAM in 5MHz bandwidth. In this thesis, the ADRES instance used for the presented analysis is an evolution of the one used for that application. Here, the ADRES features an architecture with programmable frequency domain parallel processing. However, it contains limited flexibility in time domain processing. Fig. 9.5 illustrates the 4x4 tiles of the ADRES instance under analysis. This instance contains a VLIW-section with 4 column-width-instruction. It also contains a CGA section with an array of (4x3) FUs. The FUs within the VLIW domain communicate via a multi-port global data register file whereas the FUs inside the CGA domain can communicate via global data register file, local register files and dedicated interconnects between the FUs.
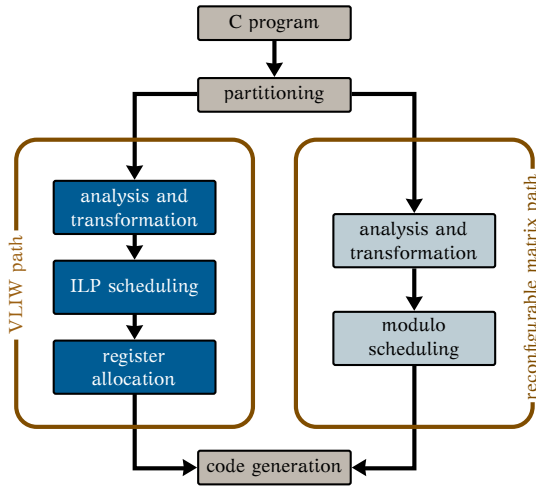
Figure 9.6.: ADRES compiler strategy.

VLIW PROCESSOR

As briefly mentioned before, the VLIW section in the ADRES architecture consists of a reconfigurable matrix with FUs in charge of computing the non-critical, often small, portion of the application, which is often irregular and full of control instructions. For that reason, Instruction Level Parallelism (ILP) is used to speed up these bottleneck tasks, by using VLIW approach it is possible to accelerate the execution of this bottleneck by a factor of 4 in comparison to what is typically done with reconfigurable matrices [68].

Reconfigurable matrices are commonly composed by parallel Reduced Instruction Set Computing (RISC), although RISC processor are considered simplistic and energy efficient, their efficiency is small compared to VLIW architectures. The VLIW helps improving overall performance by moving instruction scheduling to the compiler, an thus, reducing dramatically the burden on the actual VLIW processor.

COMPILER STRATEGY

The success of VLIW is due to their simplicity and capacity to parallelize the execution of unrelated tasks. In VLIW architectures, the instruction scheduling is shifted from the branch-control logic of traditional general purpose processors, to the compiler. Thus, the compiler is significantly more complex in VLIW architectures than in its more traditional counterparts.

The compiler strategy of the ADRES architecture consists of a C-compiler that analyzes the code, and partitions it between tasks that need to be ran in either the VLIW path or in the reconfigurable matrix path. In order to do

this, the compiler first identifies loops that can't be pipelined, since those can be efficiently executed in the CGA.

The remaining tasks that cannot be pipelined are analyzed in order to find instructions that can be ran in parallel, so that they can be allocated in the VLIW path. Finally, the compiler combines the two scheduling schemes into a single instruction set to be downloaded to the ADRES architecture. The compilation strategy is illustrated in Figure 9.6 on the facing page

10

Selective filtering in software-defined radio

The filtering overhead due to the large transmission bandwidth can be minimized by using a generic reconfigurable architecture. The principle is that the terminal does not process any sample that is not intended for it, since the control channel has to be received and decoded for all terminals within reach. The savings in term of number of data to process can be reduced to a maximum of 70% on the LTE-R8/9 compliant CC and possibly even more for CCs that only contain user data. The required sampling frequency after decoding the control channel is:

$$\mathbb{F}s_u = 12\Gamma_u\Delta f, \tag{10.1}$$

where Γ_u is the number of contiguous RBs assigned to the user u, and Δf is the LTE-A sub-carrier spacing defined by the 3GPP as 15 KHz, read more about the LTE specification in Chapter 3. Recall from Section 3.2 on page 29 that the minimum amount of RBs to be assigned to any terminal is 6. Thus, the minimum sampling rate for any user can be, in theory, as low as 1.08 MHz. However, this requires that the frequency position of the RBs intended for any specific terminals are known in advanced, moreover, this requires that the analog front-end can adapt quickly to that frequency band.

10.1. COST ANALYSIS AND SIMULATION RESULTS

Three filtering approaches are implemented and simulated in a cycle accurate simulation. Complexity comparison is done by extracting the number of clock cycle (c.c.) needed to extract the user assigned bandwidth (Γ_u) from a 100 MHz total transmission bandwidth.

10.1.1. CHANNELIZATION METHODS

The methods mapped into the CGA described in more detail in the coming subsections, in some cases the desired bandwidth is not centered at the DC carrier and a low pass filter implementation does not suffice. In order to avoid implementation of pass band filters a lower complexity digital mixer implementation is chosen. The digital mixer shifts the signal to the desired frequency so that a simple low pass filter can be applied. The above filtering methods are implemented on the ADRES platform as various software instruction sequences.

FIR-BASED SOLUTION

Consists of a FIR low-pass filter with a variable number of coefficients, this illustrated in Figure 10.1. The number of coefficients varies according to the LTE-A deployment scenario, i.e., for scenarios 8 and 5 the number of taps required to extract a single component is 23 and 102, respectively [27], read more about the aggregation scenarios in Subsection 3.2.3 on page 32.
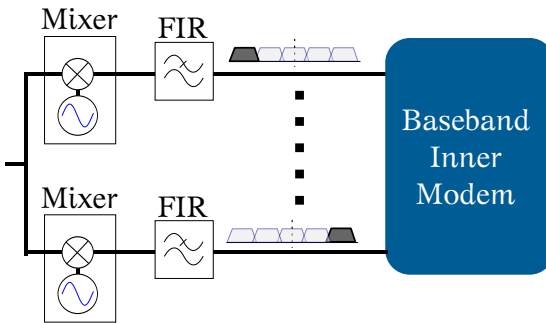
Figure 10.1.: Conceptual diagram of filtering based on a long FIR filter.

A regular FIR filter is a convolution of the incoming samples with the impulse response. Given that ADRES takes advantage of instruction parallelism, the convolution turns out to be a very expensive operation. The number of cycles needed to implement FIRs increases proportionally to the number of taps in the filter. In the current study, the minimum and maximum number of taps are 23 taps and 102 taps for scenarios 8 and 5 respectively [27]. Thus, mapping such filters in ADRES requires at least 102 clock cycles per input sample, which in comparison to 4 cycles per sample required for the FFT denote a clear interest to keep the number of taps as low as possible. A low pass filter typically requires less taps than a pass-band filter, hence, a digital mixer is placed before the filter. In this manner, the signal is shifted to the central frequency at each channel. The mixer is implemented with a special-

ized hardware accelerator that computes an angle rotation in only one clock cycle, which results in a small overhead compared to a equivalent solution with a pass band FIR filter.

10.1.2. CIC FILTER + FIR

It consists of FIR filter in combination with a decimation CIC Filter so that the number of taps of the FIR can be reduced and consequently minimize the number of samples to process in comparison to the long-FIR solution, the conceptual diagram is illustrated in Figure 10.2. For the worst case, 102 taps FIR from scenario 5, the number of taps is reduced to 34 with a decimation factor $M = 3$. Even though the CIC Filter adds some computational effort, the reduction of required clock cycles is significant (30% in this case).

It should be noted that a CIC filter provides a poor frequency response that adds noise to the pass band which is typically taken care of in posterior stages.
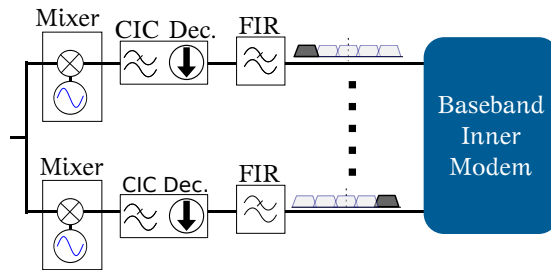
Figure 10.2.: Conceptual view of CIC Decimator + Short FIR implementation for SDR filtering.

10.1.3. FFT FILTERING

The SDR architecture can make efficient use of instruction level parallelism. The mapping of the FFT to the target architecture requires approximately 4 clock cycles per sample. Therefore, it is attractive to explore the possibility to filter directly in the frequency domain directly with the following requirements: the number of samples has to be a power of 2 and the distance between the central component carriers has to be a multiple of 300 KHz [69] i.e., the minimum multiple of the channel raster (100 KHz) and the sub-carrier spacing (15 KHz).

This is illustrated in Figure 10.3 on the next page, this is presumably the most effective way to extract a large bandwidth, however, it must be noted that this solutions is somewhat idealistic, e.g., it does assumes that either the transmission over the entire 100 MHz is done by a single BS, or else, synchronization between the CC can become a significant problem.
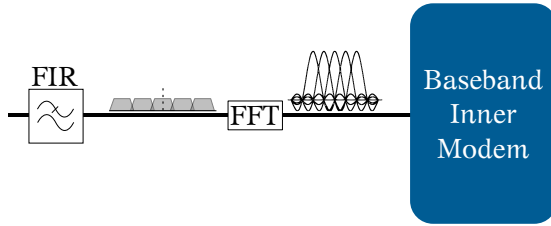
Figure 10.3.: Conceptual view of the large-FFT SDR filtering solution.

10.2. COMPLEXITY ANALYSIS

Selective filtering method was analyzed from two perspectives, cost and performance. To identify cost, the three filtering methods were implemented on the ADRES platform and its corresponding clock cycles were extracted. Energy savings are achieved by decreasing the processing time, i.e., the time the platform spends calculating the required output, which is directly proportional to the number of clock cycles used, given the assumption that the processor is operating at maximum speed, while switching to low-power mode after data processing is finalized.

Figure 10.4 on the facing page shows the number of clock cycles the SDR platform spent to filter the data with respect to the number of LTE-A resource blocks assigned to a specific terminal. The minimum bandwidth assigned to any user defined by the LTE-A standard is 6 RBs, therefore neither the clock cycles nor the RBs reach a zero value.

The method that utilizes the least number of clock cycles is the FFT, which is mainly due to architecture optimization; i.e., the architecture operates in VLIW mode taking advantage of the instruction parallelism of the FFT. Moreover, functional units capable of parallel interleaving, rotation, additions, multiplications, etc., contribute to make the FFT a very efficient implementation [70]. In contrast, the FIR turns out to be expensive and increases the required cycles proportionally to the number of taps. The dark-colored region between the FFT and the "CIC-Filter + FIR" illustrate the benefits from the optimized architecture.

The light-colored region above the method "CIC-Filter + FIR" and "Long FIR" denote the benefits of sample reduction with the aid of the CIC-Filter decimation. It can be seen that the computational efficiency is highly dependent on the correct architectural choice. Since the methods are mapped onto the same SDR platform as software instructions area comparisons are not feasible. When a full baseband implementation mapped on ADRES is considered, the percentage of clock cycles belonging to the channelization accounts for 6% for the FFT, 12% for "CIC-Filter + FIR" and 12% for the "Long FIR". Note that even though the FFT is considerably more efficient that the
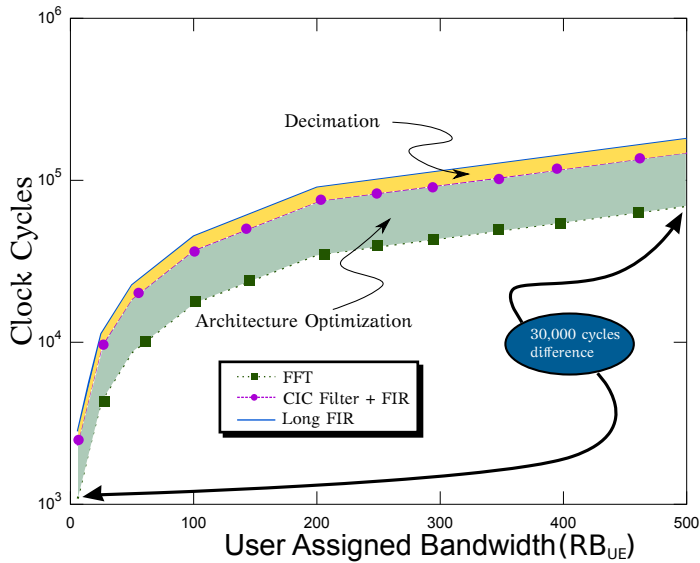
Figure 10.4.: Number of clock cycles vs. User LTE Assigned RBs (12 sub-carriers per RB), extracted from an ADRES cycle accurate platform simulator.

other methods, it is the most idealistic since it considers all CCs to have its power uniformly distributed, and perfectly aligned in frequency. Moreover, when placed into the context of the entire baseband, the difference between any two methods is at most 6% clock cycles of the entire baseband.

10.3. PERFORMANCE ANALYSIS

The performance of selective filtering was analyzed by a fully-fledged LTE Matlab simulation chain for the three typical models defined by 3GPP; EPA, EVA, and ETU, with speeds of 2, 30 and 130 Km/h respectively, for more details about 3GPP channel models consult Subsection 3.2.5 on page 35 and its references. The simulated BER is depicted in Figure 10.5 on the following page. The curves denoted as "partial" are the BER curves where only the user specific bandwidth was extracted from the entire transmission bandwidth, filtered and decoded, while the "full" BER curves denote when the entire bandwidth is filtered, decoded and only the user specific data is extracted. It is clear that the selective (partial) filtering introduces performance loss when compared to the full filtering, this is due to the fact that on the partial case, only the reference signals within the current sub-frame and within the user specific bandwidth (Γ_u) are used for the channel estimation and consequently
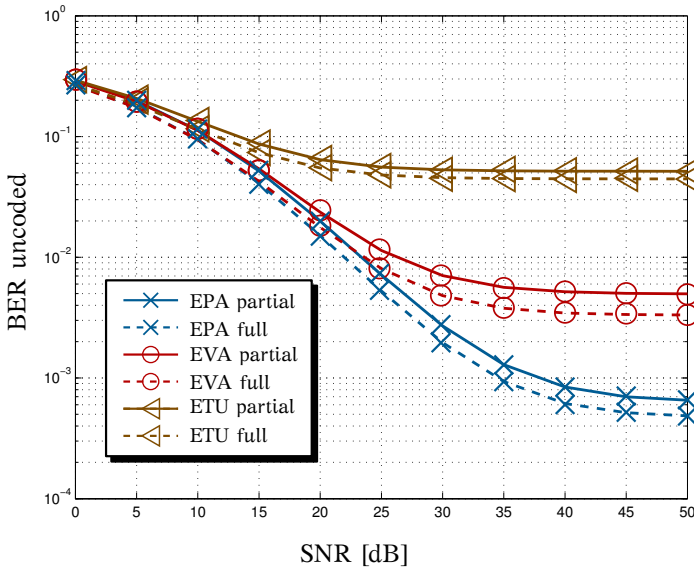
Figure 10.5.: Full precision LTE-A Matlab simulation of 100MHz
Bandwidth for 4x4 MIMO; channel estimation per-
formed by zero forcing with no coding under EPA,
EVA and ETU channel models.

for the information decoding. However, it can be seen from the figure that
the difference in terms of BER is marginal.

Consider the sample rate reduction when only the central 6 RBs are as-
signed to a specific user. A Full LTE-A with 100MHz bandwidth (500 RBs)
requires a sample rate of 5×30.72 MHz. During the first three OFDM symbols
(Control channel), since there is no previous knowledge on which CC carries
the control information, then the terminal would have to compute samples
at 153.6 MHz, then it would quickly adjust to filter the central 6 RBs at a
sample rate given by (10.1). Which according to Figure 10.4 on the previous
page results in a reduction of approximately 30 K samples for FFT filtering,
an even larger reduction for the other filtering methods. The sample reduc-
tion is crucial as the following baseband operations operate at the sample
rate generated after the channelization.

11

Summary of Part II

In the current part a feasibility study has been carried out with the purpose to simplify filtering modules inside the DFE. A simplification in these modules is beneficial when supporting multiple aggregation scenarios required by LTE-A.

The ADRES architecture is used as the target platform to implement a set of filtering methods in software instructions. The various methods are analyzed in terms of cost overhead. The cost is presented in terms of number of clock cycles required per LTE-A sub-frame. This metric is chosen as the number of clock cycles proportional to the energy consumed by given algorithm [71].

From the analysis and results it was found that the most efficient filtering method was an FFT to the entire transmission bandwidth. However, this method is also the most idealistic since assumes, among other things, perfect frequency alignment, uniform signal power, and an analog front-end capable of delivering high data-rates. Further studies are needed to modify the ADRES template so that it is optimized for other type of filtering methods.

In this part is also suggested that by using a SDR platform for filtering purposes it is possible to save up to 30K samples. The savings take place when only the minimum amount of resources is assigned to a specific user, while the LTE-A infrastructure is transmitting with a maximum aggregated bandwidth, i.e., 100 MHz.

Part III.

Multi-standard support

The current part describes the work carried out as part of the European Union (EU) project named Scalable Multi-tasking Baseband for Mobile Communications or Multibase for short. The project aim was to provide important insights on what are the trade-offs in terminal design when immerse into a multiple RATs environment. The final outcome of the project was a demonstration of a DFE on both UL and DL. Even though the author was not the sole project's contributor, the projects outcome accounts for a large portion of this dissertation. These particular contributions can be found into the following published articles:

- DIAZ, I., ZHANG, C., HOLLEVOET, L., SVENSSON, J., RODRIGUES, J., WILHELMSSON, L., OLSSSON, T., VAN DER PERRE, L., AND ÖWALL, V. "A New Digital Front-End For Flexible Reception In Software Defined Radio", *Elsevier, Microprocessors and Microsystems: Embedded Hardware Design, invited to special issue on NORCHIP*, March, 2014.

- DIAZ, I., ZHANG, C., HOLLEVOET, L., SVENSSON, J., RODRIGUES, J., WILHELMSSON, L., OLSSSON, T., VAN DER PERRE, L., AND ÖWALL, V. "Next generation digital front-end for multi-standard concurrent reception", *NORCHIP Conference*, November, 2013.

- ZHANG, C., DIAZ, I., ANDERSSON, P., RODRIGUES, J., AND ÖWALL, V., "Reconfigurable cell array for concurrent support of multiple radio standards by flexible mapping," *IEEE International Symposium on Circuits and Systems*, May, 2011.

- ZHANG, C., DIAZ, I., ANDERSSON, P., RODRIGUES, J., AND ÖWALL, V. "Reconfigurable cell array as enabler for supporting concurrent multiple standards in mobile terminals," *10th Swedish System-on-Chip Conference*, May, 2010.

12

EU multibase project

There is an increasing number of communication standards[1] created, despite the efforts of standardization bodies and governments to keep the total number limited. A RAT is typically optimized for a specific set of features or for an application, e.g., data-rate, power consumption, coverage, etc. Since some of these features derive contradicting implementations, it is difficult to imagine that a handful of RATs will ever dominate communication over radio waves.

Thus, future mobile handsets will need to concurrently support a variety of RATs, as mobile terminals require connecting to specific services via different interfaces. A key barrier for the support of end-to-end connectivity lies in the difficulty to implement cost-efficient mobile terminals that simultaneously support various blends of these RATs. Different RATs provide different data-rates under different mobilities. The future mobile terminal will be able to move from RAT to RAT. This terminal should dynamically adapt its internal structure in accordance to the instantaneous mobility and data rate required. In this manner, vertical hand-over could give users the freedom to move from one RAT to another in an optimized way that directly impacts the service perception. This optimization factor can be service cost, quality, performance, etc.

Simultaneous support in either Tx or Rx is still a big challenge in both analog and digital domains, several attempts to address this issue has been reported in both industry and academia [72], [73], and [74]. However, to this day, this is still a question of research with no definitive answer. Multibase project was created to address some of these issues and provide insights to those questions.

---

[1] Note that, the words standard and Radio Access Technology (RAT) are used interchangeably across this document.

12.1. DESCRIPTION

Multibase project was created as a combination of European leading partners in industry and academia in order to strengthen Europe's leading position in high-speed, end-to-end, mobile systems technology. The team incorporates three world-ranked universities (Linköping, KU Leuven, and Lund Universities), two multinational companies (Ericsson and Infineon), a leading international research center in nano-electronics IMEC, and a well recognized technology based consultant (Technikon). A brief description of each partner is listed as follows:

PROJECT PARTNERS

- **Technikon - Austria**: Privately owned research company in the microelectronic sector with a highly specialized multinational team of scientists and engineers. Acts as project coordinator, handles the project and consortium management. It is also responsible for reporting, controlling, assessment and dissemination of the project progress [75].

- **Ericsson - Sweden**: World-leading provider of communications technology and services, with more than 110,000 people working with customers in more than 180 countries. By 2013, 40 percent of the world's mobile traffic goes through Ericsson networks, supporting customer's networks service to more than 2.5 billion subscriptions [76].

- **Infineon Technologies - Austria**: Infineon Austria develops and produces microchips for applications in the area of mobility, industry and security. With employees from some 45 countries engaged in the research, development and production of microchips at our sites in Villach, Klagenfurt, Graz, Linz and Vienna [77].

- **Lund University - Sweden**: International university with global recruitment founded in 1666, with 47,000 students and 7,200 employees based in Lund, Malmö and Helsingborg. The University has a turnover of around EUR 750 million, of which two thirds is in research and one third in education [78].

- **Linköping University - Sweden**: Is one of Sweden's largest academic institutions, with research activities in a variety of disciplines. With its 3,900 employees, it host 1,400 research students, 27,000 undergraduate students in 14 multidisciplinary departments [79].

- **IMEC - Belgium**: It is a world-leading research center in nano elec- tronics with headquarters in Leuven, Belgium. With state-of-the-art technology and tight connections with industry and academia, IMEC contributes with the publication of a large number of scientific papers, various patent applications, and several spin-off companies [80].

- **KU Leuven - Belgium**: Founded in 1425 by Pope Martin V, KU Leu- ven bears the dual honor of being the oldest extant Catholic university in the world and the oldest university in the Low Countries, it employs 10,255 people, and hosts around 41,000 students per academic year, it has extensive activities in research of various disciplines and is con- sidered one of the leading universities for technology development in Europe [81].

PROJECT GOALS

The RAT selection was done based on OFDM, which is predicted to be one of the dominant access technology in the next decade. Thus, supporting concurrent OFDM standards is a crucial requirement. Even though there is some research on reconfigurable hardware for multiple RATs [72], no attempt has been done in focusing on OFDM-standards alone (taking advantage of their similarities), nor in considering concurrent support.

  As a proof of concept, it was decided to focus on three OFDM RATs that provide complimentary services, namely, WLAN contributing to high data-rate under stationary conditions, LTE for high mobility with moderate data-rate and DVB-H providing multimedia broadcasting services. Since the selected standards are OFDM, reuse in several hardware blocks is possible, while differences make reuse non-trivial. These differences and similarities are described in more detail in Chapter 3.

  It is foreseen that SDR will be a key element on providing the required flexibility in future baseband processors, however, there is little research re- garding the DFE capable of providing enough data to be used by the baseband processor in question. Thus, the architectural focus of the project is placed in the DFE in both UL and DL and its interconnection to a more general SDR-based baseband processor.

  The ultimate goal is to demonstrate the feasibility of this architecture with today's technology, so that it can serve as ground basis for the developers of mobile terminals in the coming decade. The extent to which this can be done with the resources available for the project is discussed in Section 12.3 on page 112.

12.2. ORGANIZATION

The team selected three areas where the focus of the activities was placed, namely:

- Multi-tasking radio,

- scalable reconfigurable multi processor technology, and

- algorithm/architecture co-design for maximum energy efficiency.

The project structure and its corresponding deliverables are shaped in accordance to the aforementioned areas, with intention of demonstrating innovative ways of realizing architectures for baseband terminals. Work Packages (WPs) are created as the baseline of the project where the intended objectives are broken down into individual WPs. Partners participate in several WPs according to its corresponding competence and expertise. The six WPs are described as follows and its interaction can be graphically visualized in Figure 12.1 on the next page.

WP01 - SYSTEM ENVIRONMENT AND USER SCENARIOS

Knowledge of the future in wireless communication is crucial for the role of this WP, important parameters such as minimum data-rates, maximum transmission power, and energy budget in the terminal, are required to select the possible applications likely to be dominant in the next 10-15 years. WP01 does not concentrate on state-of-the-art technology, but rather looks beyond it and understand the system and terminals trade-offs involved in the process. It must also take into consideration the business part of the picture, so that, the platform functionality is optimized for best user experience at the same time as being being good monetary incentive to handsets designers. The main contributor and leader of this WP is Ericsson.

WP02 - ARCHITECTURE ALGORITHM CO-DESIGN

This WP focuses on optimization taking into consideration "performance-complexity scalable design paradigm". As briefly discussed in Chapter 1, ASICs are typically energy efficient with fixed functionality, yet they lack flexibility and scalability for an adaptive architecture. Reconfigurable architectures can prove a better alternative if the structural complexity is correctly defined. The external specification previously defined by WP01 is taken as input to this WP, where the system environment, the user scenarios, and the definition of the platform framework are considered in order to design the best possible algorithms, see WPs relation in Figure 12.1 on the facing page. This in turn, results in the best use of hardware resources while taking a cross-disciplinary approach between algorithm and architecture. The work in this WP is mainly led by IMEC.
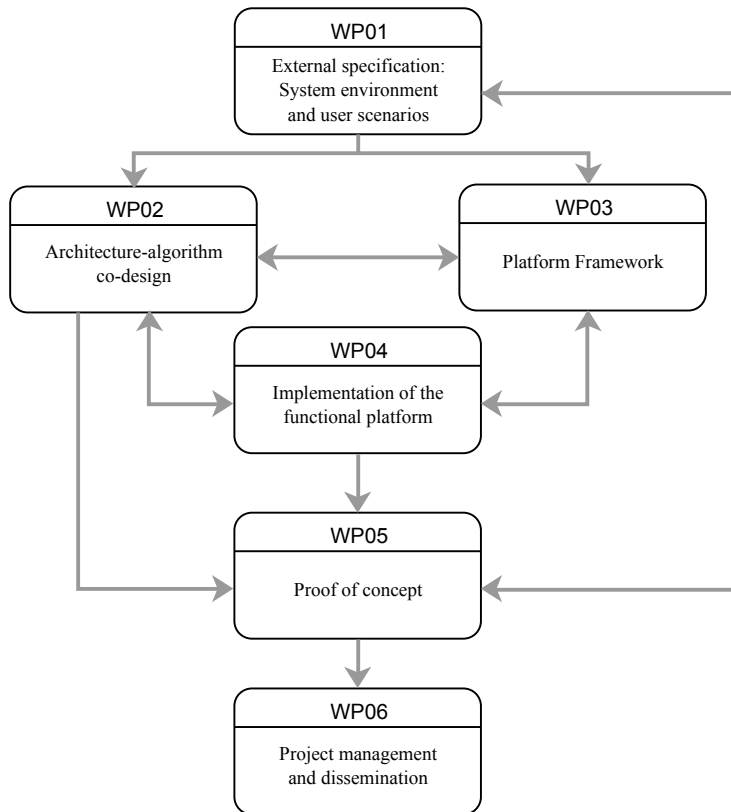
Figure 12.1.: WP interaction in the Multibase project.

WP03 - PLATFORM FRAMEWORK

Given the complexity of today's modern mobile terminals, an efficient way to communicate between various parts of the system is essential. In this work package, the activities are focuses on studying different SoC communications alternatives, high-speed buses as well as NoC or a combination of the two are needed to provide ultra low (connection and data) latencies. Moreover, it is also crucial in this WP to evaluate specific supporting alternatives to heterogeneous structures and variety of data types. Another aspect to consider is the appropriate memory scheme needed for maximum flexibility, e.g., memory addressing type, level programming supported, etc. Linköping University is appointed the main contributor to this WP.

WP04 - IMPLEMENTATION OF THE FUNCTIONAL PLATFORM

The main target of this WP is to develop and implement the architecture capable of addressing the challenges imposed by multi-RAT support, multi-tasking of several simultaneous independent connections is possible with separate dedicated hardware modules, however this solution is expensive, inefficient and consequently undesirable. Possibly a MPSOC architecture can handle the large complexity while assuring moderate cost in silicon and the targeted flexibility. Thus, the main tasks here are: 1) to develop the overall functional architecture with respect to granularity of the processing architecture and the physical inter communication structure based on WP03 results with special emphasis on multi-tasking radio, and 2) to design and implement key architectural building blocks identified in WP02 with the technology perspective from WP05. Lund University is the partner in charge of supervising and carrying out the work done in this WP.

WP05 - PROOF OF CONCEPT

The actual realization of a prototype is essential in demonstrating the achievements in the project, therefore this package investigates challenges in sub-micron technology, e.g., temperature, systematic interconnect, and process variations due to lithography. In addition this WP provides the means towards an actual proof of concept fabrication of critical blocks that give sufficient validation of the Multi-Base platform. The partner appointed to be in charge of this WP is Infineon.

WP06 - PROJECT ADMINISTRATION

The main task of this WP is to continuously monitor the project's health in terms of goals and budget. In addition, it also provided the IT-infrastructure to make the interaction between different WPs and partners as fluid as possible. The main partner in charge of this WP is Technikon.

12.3. ACHIEVEMENTS

Given the limited size of the project, a complete terminal demonstration was not within reach. Instead, three elements were found crucial for the well-functioning of the terminal. These elements were chosen as the main subject to optimization. The remaining parts are analyzed theoretically and simulated. The project's final architecture and the interaction between its various modules is depicted in Figure 12.2 on the next page.

The elements selected for fabrication are related to transmission, reception, and interconnection, namely, the Digital Front-End Transmitter (DFE-Tx), Digital Front-End Receiver (DFE-Rx), and platform controller, the remaining parts in the architecture were also considered, but not taken down to silicon

Multibase platform
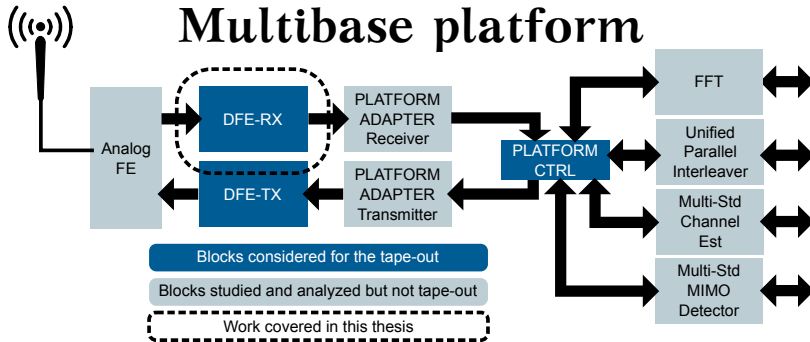
Figure **12**.2.: Block diagram of the Multibase conceptual terminal, three main crucial components (DFE-Rx, DFE-Tx, and platform controller) are identified, fabricated and verified during the project.

implementation. More details about the analysis on those parts can be obtain via the public project deliverables, e.g., the multi-RAT channel estimation and FFT are analyzed on algorithmic level in [82], whereas the description of the unified parallel interleaver and the multi-RAT mimo-detector can be found in [83].

The project success was achieved when the three fabricated modules where successfully demonstrated to EU officials, the final outcome of the project is a comprehensive examination of limitations and capabilities of today's technologies for multi-RAT transmission and reception, all available in public reports of the EU commission.

The author contributed to design, development, fabrication and demonstration of the DFE-Rx, particularly within the activities of WP04 where Lund University was in charge. Thus, the functional architecture constitutes an important part of this thesis, the author also contributed by developing the synchronization algorithm used in the final architecture. The subsequent chapters within the current part are specifically related to the analysis done while working on the DFE-Rx, e.g., the functional blocks and its algorithm are detailed in Chapter 13, while the fabrication and its testing is described in Chapter 14.

13

Functional platform

A brief description of each of the standards targeted is presented in Chapter 3, furthermore a special section with their similarities can be found in Section 3.5 on page 43. In this chapter those similarities are employed as means to reduce complexity, this requires close interaction between baseband-algorithm and implementation.

Since the main goal is to support at least two standards, the possibility of reusing components adds an extra dimension to the usual performance-complexity trade-off when considering a single standard. The work is then divided into parts of the receiver, rather than into modules to individual RAT.

13.1. SPECIFIC REQUIREMENTS PER STANDARD

An analysis on the distinct features of each RAT needs to be done in order to account for the appropriate adjustments in the functional architecture. Specifically those affecting the DFE functionality, e.g., sampling rate, synchronization specific structure, bandwidth, mobility, etc. A few of them are addressed as follows.

13.1.1. SAMPLE RATE

One of the fundamental requirements of any standard is its sampling rate. The sample rate is directly associated with its transmission bandwidth, and in OFDM RATs related also with its sub-carrier spacing (read more about this relation in Subsection 2.2.3 on page 18). Since the DFE is placed directly after the Analog to Digital Converter (ADC), it is needed parts of the architecture run at sample rate; i.e., the sample rate is equal to the clock frequency. However, since the elementary sampling frequency is different in

each standard.

In such case there are two alternatives to consider:

- The first consist of operating parts of the DFE at different clock rates each matching the specific standard to be received. This is, however, not an attractive choice since the sampling rate of the intended RAT is not an integer multiple of one another, thus, complex clock synchronization circuits such as multiple Phase-Locked Loops (PLLs) would be needed.

- The second alternative, the preferred choice, consist of operating the entire architecture at a single speed that matches the sample-rate of the fastest standard. While re-sampling the other(s) so that orthogonality is preserved. For that re-sampling methodologies need to be explored. This will become clear when the decimation chain is discussed in Section 13.2.3 on page 118.

13.1.2. SYNCHRONIZATION

From the DL perspective, there is a fundamental difference between WLAN and the other two standard specifications. WLAN communication is done by burst transmissions, contrary to what happens in LTE and DVB-H. For LTE and DVB-H the base-station is transmitting synchronization information periodically, whereas WLAN provides a preamble for purposes of acquisition (see WLAN specification in Section 3.3 on page 37). In order to have the maximum possible flexibility in the architecture, it was decided to have a programmable CGRA that would be capable of loading a specific algorithm on demand. Hence, it is possible play on the fly with variables such as: resource management, computation accuracy, and latency.

The following section details the architecture defined for this project and each of its components.

13.2. ARCHITECTURE FUNCTIONALITY

The ultimate goal defined by the project is to support two or more RATs concurrently. Thus, different modules contribute with different levels of configuration, e.g., control, filtering and down-sampling, are very specific to each individual standard and need to be operational during the entire transmission, while the synchronization engine is required only during the first stage of connectivity. Thus, the proposed DFE needs a highly reconfigurable data-path per standard and a single synchronization engine linking each data-path. This architecture is graphically illustrated in Figure 13.1 on the facing page, where a maximum of two concurrent standards (LTE, DVB-H, or WLAN) are supported concurrently, however, since not 100% of the architecture is
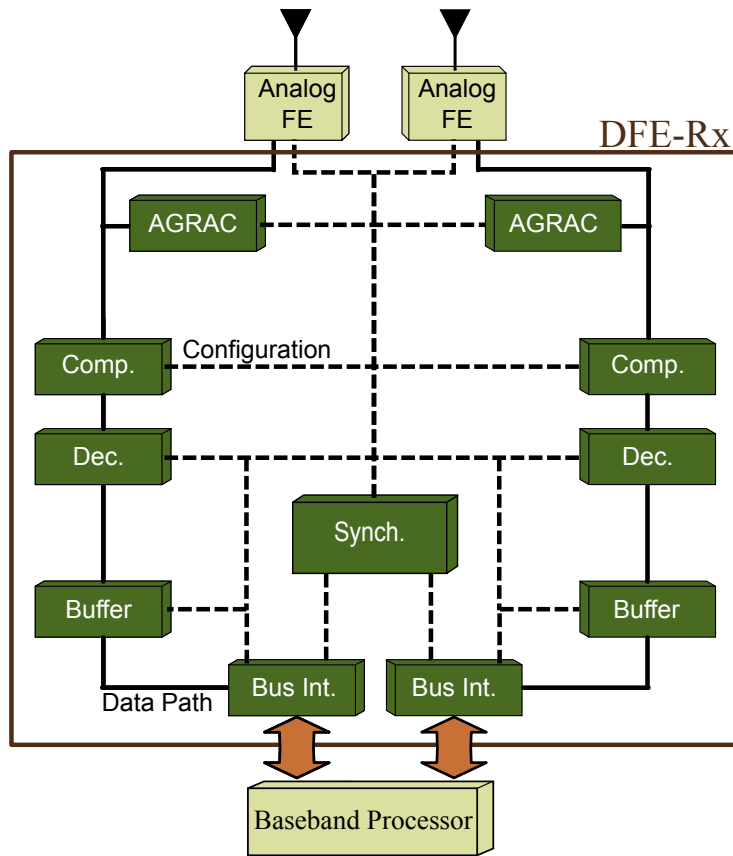
Figure 13.1.: DFE-Rx top level.

constantly used, more concurrent standards could be supported. The architectural building blocks are detailed in the following subsections.

13.2.1. AUTOMATIC GAIN CONTROL

The Automatic Gain and Resource Activity Controller (AGRAC) serves two major tasks. The first is to control certain front-end settings that require a short control loop, such as front-end gain and DC-offset compensation. The second task is to control which other parts of the DFE that are active at a certain moment in time. Hence, it also controls the mode of operation of the DFE. Depending on the standard to be initiated, it will enable filters, receive data buffer and synchronization engine. When the system has synchronized with the recently detected standard, the AGRAC generates the interrupt signal towards the baseband processor.

The AGRAC is implemented as a Micro-Controller (µC), compatible with the architecture of a Microchip PIC16F84A controller [84]. Selecting an architecture that is compliant with an industry-standard has the advantage that the existing tool-chains for application development and debugging can be extensively used.

13.2.2. COMPENSATION

This block performs a compensation for DC-offset and IQ imbalance compensation. The amount of offset present in the incoming signals is estimated and communicated to the compensation module by the AGRAC. The compensation module corrects the error by adjusting the values of the in-phase and quadrature components accordingly.

13.2.3. DECIMATION CHAIN

A multi-standard receiver must obtain the baseband signal at a sample rate given by the received standard. The master clock has a frequency unrelated to the standard's sample/symbol rate, and is sufficiently high to capture the signal content without aliasing. The decimation chain brings the signal to correct sampling frequency in the digital domain. This is done by filtering and down-sampling to the specific standard baseband sample-rate. By adjusting parameters inside the decimation chain, it is possible to digitally change the sample rate of the received signal to the required for each supported standard.

The decimation chain block consists of programmable filters that remove unwanted signals and keep only the desired signal, and a Farrow resampler can adjust the sample rate of the signal by an irrational factor [85], this factor depends strongly on the configuration of which standards are being received. The block diagram of the decimation chain with its main components is illustrated in Figure 13.2 on the next page.

The flexible filter consists of a symmetric FIR filter with 32 taps, the architecture is folded in order to reduce the number of multipliers to 4, with coefficients dynamically changed in dependence of the required standard configuration. The Farrow resampler achieves the sampling transformation by a combination of filter banks and polynomial evaluation, where polynomial and filter coefficients are manipulated in relation to the current standard configuration.

13.2.4. SYNCHRONIZATION ENGINE

The baseline architecture for implementing the multi-standard synchronization is based on a dynamically reconfigurable CGRA [86]. The reconfigurable array is constructed from a mesh of heterogeneous resource cells, communicating through a combination of local interconnections with dedicated wires and a global hierarchical routing network. To meet different computation de-
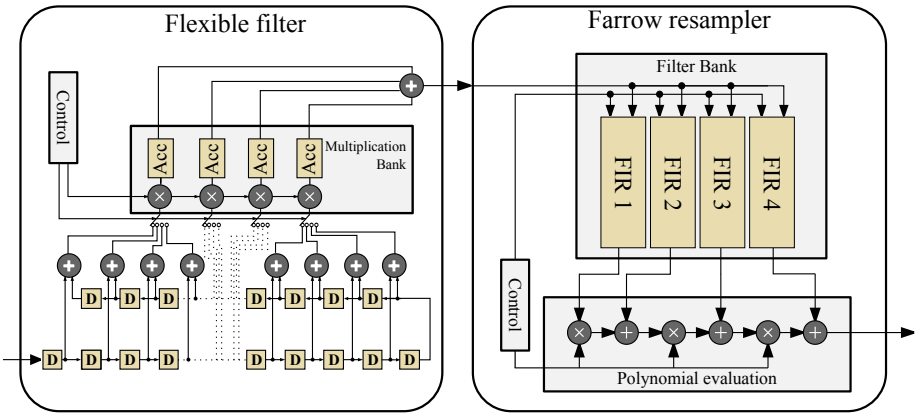
Figure 13.2.: Block diagram of the decimation chain consisting of a flexible filter and a farrow resampler. The filter is a 32 taps symmetric FIR filter with configurable impulse response. The farrow resampler contains configurable coefficients filter bank and dynamic polynomial evaluation.

mands, the flexible hardware architecture allows the user to allocate different function blocks in cell array, comprising both application-tuned or general-purpose processors and data memories. Figure 13.3 on the following page illustrates the structure of a) processor cells, c) memory cells, and b) the interconnection between resource and memory cells in the CGRA architecture.

Processor cells implement both arithmetic and logic operations to map applications to the CGRA, while memory cells act as storage buffers for both data processing and communication. Each memory cell contains one or more memory banks, where each bank can be configured to emulate either FIFO functionality or to support Random Access Memory (RAM). All resources can be parameterized at system compile-time, and are capable to be dynamically reconfigured to support run-time application mapping. A more detailed description of this architecture is published in [87].

In order to perform time synchronization on two concurrent standards, it is required to buffer data from both standards. The buffering can be accomplished by operating memory cells in FIFO mode. Multiple memory cells operating in FIFO mode can be concatenated at system run-time to form a large data storage, providing the user more flexibility on memory usage. This feature fits nicely in multi-standard applications, as different standards often have different requirements on memory capacity. Using a large data memory to cope with all standards under consideration, results in poor memory utilization. Hence, partitioning the memory cell into smaller pieces increases both efficiency and flexibility of the hardware usage, as multiple memory ac-
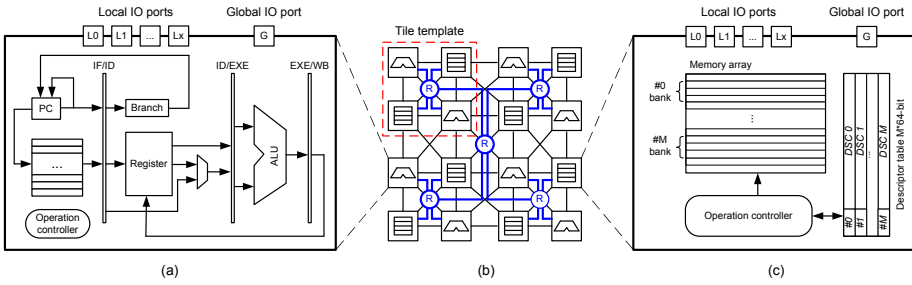
Figure 13.3.: (a) Building blocks in processor cell. (b) The CGRA is constructed from an array of processing and memory cells, connected using a local and a global hierarchical routing network. (c) Architectural block diagram of a memory cell.

cesses are provided simultaneously to different resource cells in the array. With the use of memory concatenation ability, applications with larger memory requirement are also supported in this CGRA. The provided memory flexibility is applied to the three radio standards considered in this study.

13.2.5. BUFFER AND BUS INTERFACE

The DFE-Rx communicates with an SDR baseband processor with high speed communication interfaces. Thus, the DFE-Rx needs an advanced bus interface to be able to stream the data processed by the DFE-Rx that is temporarily stored in the buffer. The need for this buffer arises from the fact that WLAN communicates in bursts. Hence, every sample contains relevant unrecoverable information during the initial connection stages.

14

DFE-Rx proof of concept

In order to measure the real benefits of the proposed architecture in, the actual implementation of the functional architecture was carried out; in this manner the costs and benefits can be measured in terms of silicon area and power dissipation. The current chapter presents such numbers in addition to introducing the verification strategy used in the process of measuring.

14.1. SYNTHESIS OF THE ARCHITECTURE

The DFE-Rx was fabricated in a $65\,\mathrm{nm}$ CMOS technology with a die size of $5\mathrm{mm}^2$, Figure 14.1 on the next page depicts the final layout with its corresponding dimensions. The memories are distributed around the entire design and the pad ring is made out of 144 input and output pads.

14.1.1. AREA BREAKDOWN

The prototype design is pad-limited where according to synthesis results, $2.6\,\mathrm{mm}^2$ are used for memories, logic-related cells and pads, while the remaining area is used for power and signal routing. Tables 14.1 to 14.3 on pages 122–123 show the area break-down of the design after synthesis. Looking into the top level hierarchy in Table 14.1 on the following page, it is seen that the pads account for 40% of the entire area, and the remaining is almost uniformly distributed among the synchronization engine and the two standard's data-paths with 20% share each. Note that the large area share is expected, since it is a prototype circuit and, thus, it has a large number of input and outputs (144 in total).

Looking into each data path area distribution, it can be seen from Table 14.2 on page 123 that the largest share of the data-path is occupied by the AGRAC with 40% share (internal memories included), which is no sur-
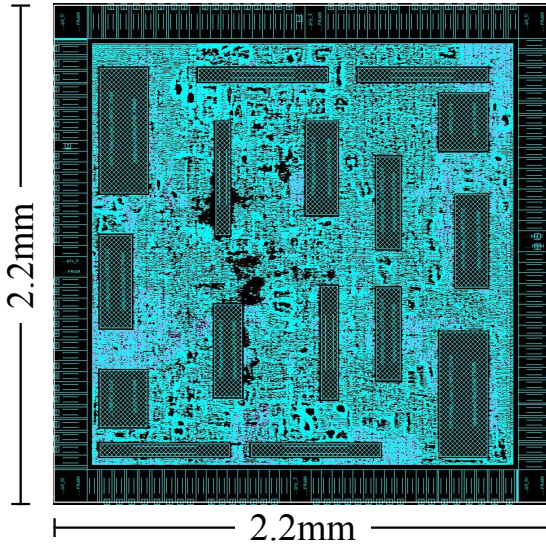
Figure 14.1.: Final Layout of the entire DFE-Rx, with 144 pads occupying a die area of $5\,\text{mm}^2$.

prise given that this module is a functional μC with all the flexibility offered by its architecture, in addition, it requires a module for programming the μC. Buffers and the host interface have the second largest area share in this hierarchical level with 27%, due to the large memories acting as buffers. In addition the host interface is capable of performing Direct Memory Access (DMA) operations to the baseband processor, requiring relatively large number of resources.

Table 14.1.: Area occupied by a single DFE-RX.

DFE Rx	Area [μm^2]	Share [%]
Input/Output Pads	991,569	38.37
Synchronization Engine	479,026	18.54
Data-Path Standard 1	501,894	19.42
Data-Path Standard 2	501,894	19.42
Data-bridge	50,676	1.96
Others	59,175	2.29
TOTAL	2,584,234	100.00

The Table 14.3 on the facing page shows the area break-down of the synchronization engine, where most is occupied by memories with a total share of 42% of the area, i.e., 13.78%, 12.04%, 7.73%, and 8.63% for the memories of memory cell 0, memory cell 1, processing cell 0, and processing cell 1, re-

spectively. The second largest contributor is an interface controller, in charge of communicating with the remaining parts of the DFE-Rx, with 12.59% area share.

Table **14.2.**: Area occupied by the Data Path for a single standard path excluding synchronization engine.

Single Data Path	Area [μm^2]	Share [%]
AGRAC	204,572	40.76
Compensation	6,034	1.20
Resampler	69,333	13.81
Buffer & Host Int	137,623	27.42
Debug Module	75,774	15.10
Others	8,557	1.70
TOTAL	501,894	100.00

Table **14.3.**: Area occupied by the synchronization engine based on CGRA for reception of one or two concurrent standards.

Synchronization Engine	Area [μm^2]	Share [%]
Router control	18,968	3.96
Router data	19,008	3.97
Memory Cell 0	39,167	8.18
Memory Cell 0 RAM	66,016	13.78
Memory Cell 1	39,169	8.18
Memory Cell 1 RAM	57,657	12.04
Internal Comm Buffer	25,640	5.35
Processing Cell 0	37,567	7.84
Processing Cell 0 RAM	37,009	7.73
Processing Cell 1	37,201	7.77
Processing Cell 1 RAM	41,318	8.63
Interface controller	60,320	12.59
TOTAL	479,040	100.00

The area share is also visually represented by the three level pie chart in Figure 14.2 on page 127, where the partition of the data-path 1 is omitted since it is the same as of data-path-2 already depicted.

14.2. VERIFICATION

Two verification strategies were developed to test the DFE-Rx as a fully-functional system and to test the synchronization engine as standalone. For these two strategies, two verification boards were created. Figure 14.3 on page 128 shows the system verification board, where the DFE-Rx is connected to an FPGA and further interfaced via another FPGA to a computer for data visualization. In order to simplify the verification of the entire platform, the computer is also capable of accessing specific debugging interfaces on the fly; in this manner it can communicate with several internal modules without disturbing its main functionality. The second board is shown in Figure 14.4 on page 128 where the synchronization engine is directly accessed by bypassing other elements inside the DFE-Rx and a standalone test is performed via a serial interface running at 10 MHz.

14.3. RESULTS

Maximum flexibility is achieved by integrating a combination of mixed-grained architectures. From highly efficient hardware accelerators, to a flexible reconfigurable CGRA platform.

When two concurrent standards are being received the synchronization engine adjusts its memory and processor cells, in a way that the hardware is split by two, and shared by the two streams by either time sharing or memory sharing. When only one standard is being received, the synchronization engine adjusts itself to exploit maximum accuracy by decreasing quantization noise and furthermore improving algorithm performance, given that the synchronization engine can store several algorithms in its program memory, it is possible to switch between them whenever needed to increase performance, decrease power consumption, etc.

In terms of hardware complexity, the high reconfiguration evidently increases complexity. In order to measure this complexity overhead, the synchronization engine has been synthesized in a 65 nm CMOS process and compared to an equivalent solution provided with only hardware accelerators. The synthesis result of the accelerator based synchronization is shown in Table 14.4 on the facing page.

By examining tables Table 14.3 on the previous page and Table 14.4 on the facing page, it is not possible to compare module by module, since the CGRA has multiple functionality. However, it is evident that both design are memory dominant, i.e., correlation FIFO and dual moving sum in the case of the accelerator and the all the RAM elements in the CGRA solution. In terms of functionality, two accelerator-solutions are equivalent to a single CGRA-solution, then it can be seen that for a reception of two concurrent standards the area overhead is a factor 2.7, i.e., the CGRA is almost 3 times

Table 14.4.: Area occupied by synchronization engine made out of
hardware accelerators for a single standard reception.

Synch Accelerator	Area [μm²]	Share [%]
Correlation FIFO	27,232	44.69
Config module	42	0.07
Complex multiplier	1,296	2.13
Dual moving sum	25,258	41.45
Peak detector	2,305	3.78
Others	1,479	2.43
CORDIC Time-Multiplexed	3,330	5.46
TOTAL	60,942	100

larger than its accelerator counterpart. However this overhead is reduced to only 1.2 when considering the entire DFE-Rx.

Figure 14.5 on page 129 illustrates the hypothetical scenario where the maximum number of concurrent standards is 32. Even though this is scenario is unpractical, serves as reference to find the point where the overhead is no longer too costly, this point is reached when a minimum of 12 concurrent standards are needed. Also from Figure 14.5 on page 129 it can be seen that in the system perspective, the overhead does not vary much by increasing the number of concurrent standards, since the area-dominant parts of the DFE-Rx have to be duplicated per standard, i.e., all the elements in Table 14.2 on page 123.

Table 14.5.: Sync Engine Maximum speeds after synthesis and number of standards to be capable of supporting if running at that speed.

Tech. [nm]	CA Max Freq. [MHz]	LTE 20MHz	WLAN 20MHz	WLAN 40 MHz	DVB-H 2K Mode
130	267*	2	3	1	7
90	385*	3	4	2	10
65	534	4	6	3	14
40	867*	7	10	5	23
28	1239*	10	15	7	33

*Scaled from 65 nm with scaling rules: $A \sim 1/s^2$ and $T \sim (1/s)$.

Moreover, the CGRA solution can be more attractive when considering technology scaling, this is depicted in Table 14.5, where the maximum speed extracted from synthesis has been scaled to various CMOS technologies. Hypothetically assuming that the fabricated synchronization engine can be clocked at this maximum speed, the number of concurrent standards that the synchronization engine can support (given each standard requirements)

is shown in the table. Even though the application of 33 concurrent DVB-H receiver signals is hard to imagine, the scalability potential provided by the CGRA is evident.

At the time of writing, system-functionality verification is still pending. However, the synchronization engine standalone verification provides a clear indication on the concept's feasibility. Measurements were performed on the synchronization engine receiving WLAN standard with a core supply of 1.2 V. The measurements is depicted in Figure 14.6 on page 129 where four consecutive patterns of WLAN are input to the synchronization engine, in the figure these are labeled by circled numbers, followed by a idle period. It is worth mentioning that during the configuration loading, there is a power consumption of 1.65 mW (Configuration phase is not shown in the figure).
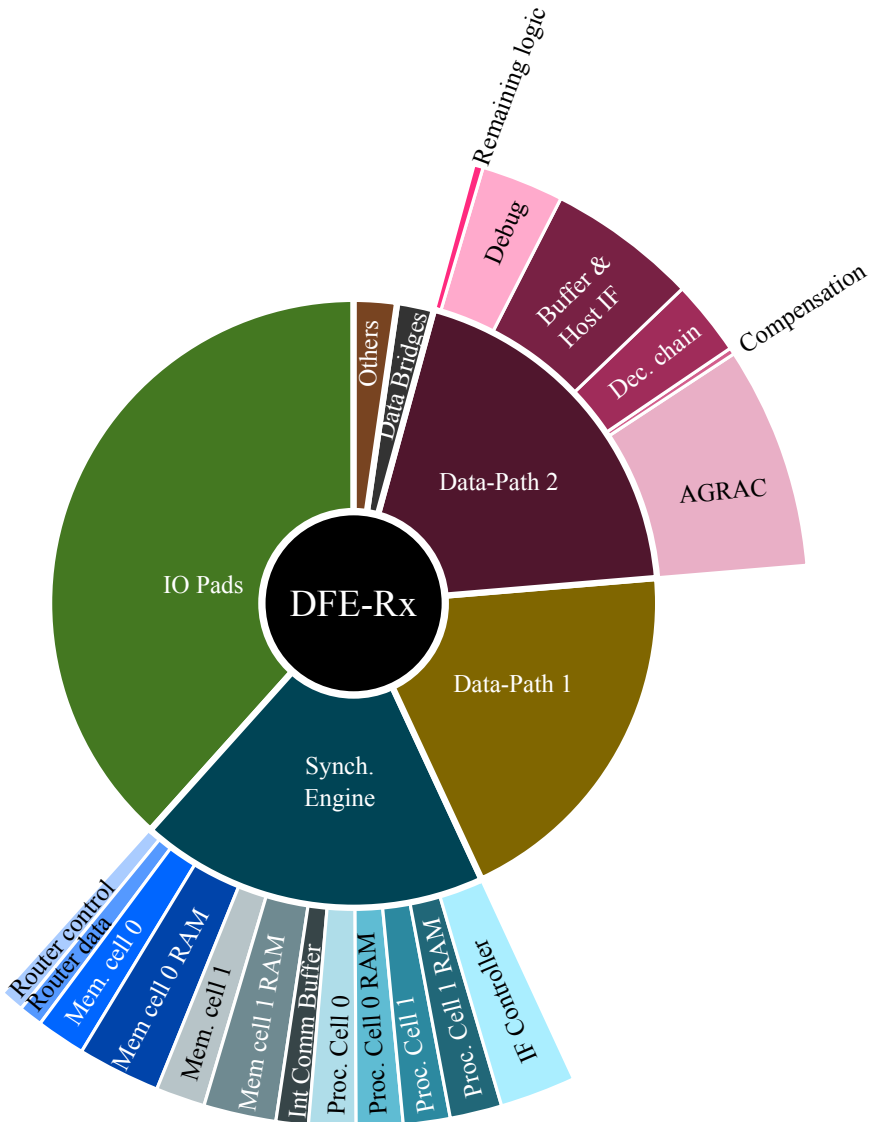
Figure 14.2.: Graphical visualization of the area share of the various parts of DFE-Rx, where the size of each element is proportional to its corresponding occupied area in silicon.
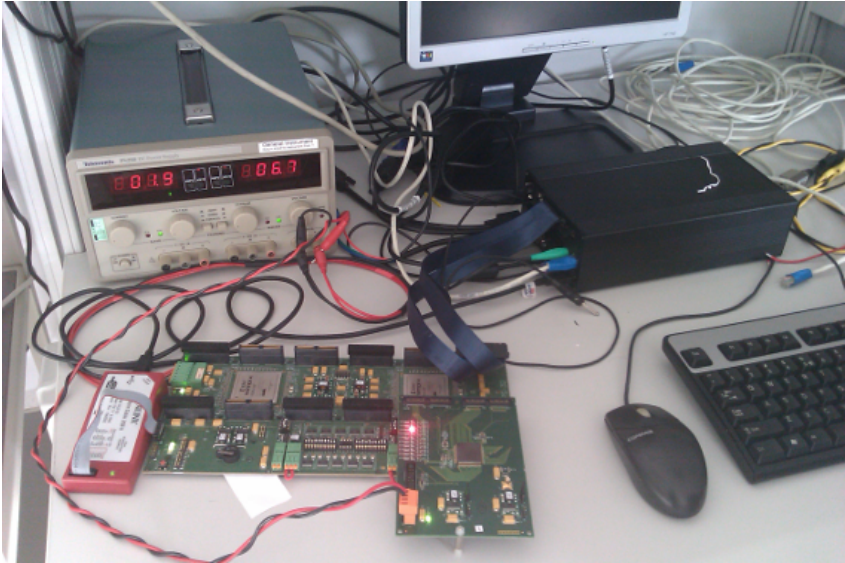
Figure 14.3.: DFE-Rx system verification setup, connected to FPGA and computer via PCI interface.
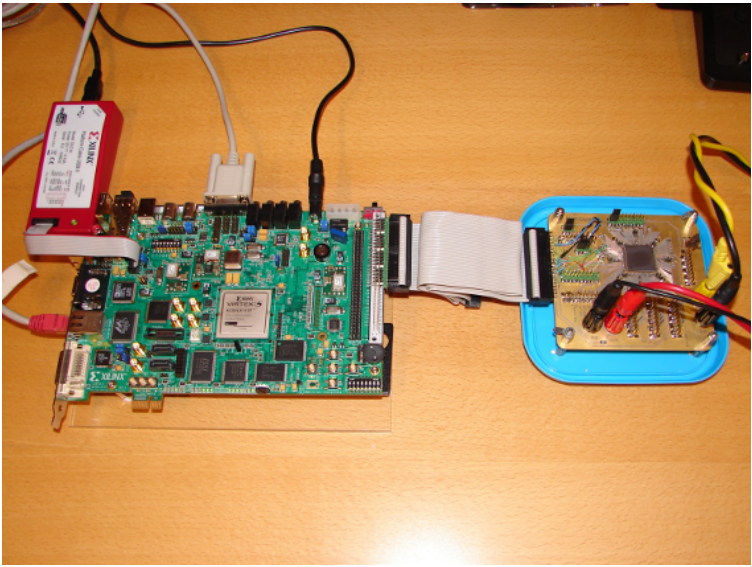
Figure 14.4.: Synchronization engine stand-alone verification setup, communication is done via serial interface under lower operational frequency and communicated to computer via serial port.
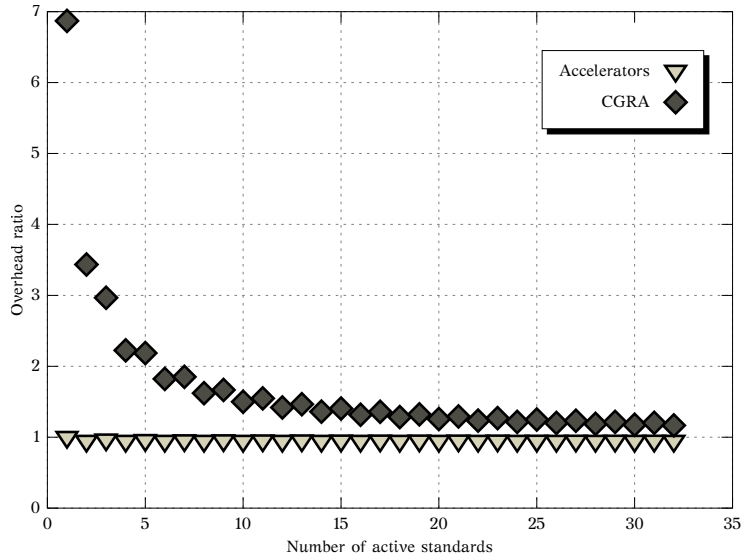
Figure 14.5.: Area Overhead of synchronization engine solution over accelerator solution shown for synchronization engine only, and full DFE-Rx.
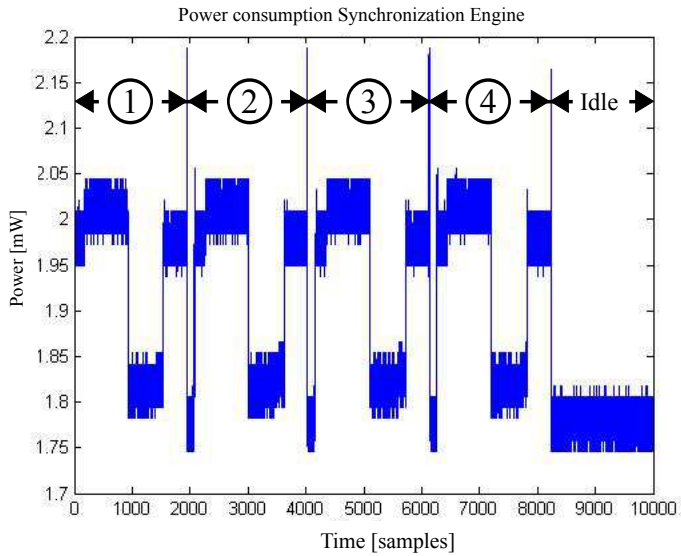
Figure 14.6.: Power consumption of Synchronization Engine as standalone when processing an incoming WLAN standard.

15

Summary of Part III

When a mobile terminal is subject to support multiple-streams concurrently, the critical points becomes an architecture that has good flexibility with limited overhead, algorithm-architecture co-design plays a critical role in the design of future terminals since the level of flexibility required by the mobiles of the future, typical hardware accelerators, Digital Signal Processors (DSPs) or simple General Purpose Processors (GPPs) are not sufficient as stand alone solutions.

In order to overcome this challenge in the current part a heterogeneous architecture is presented capable of synchronizing up to two completely concurrent OFDM standards. The feasibility of this study is shown by fabricating a DFE capable of supporting up to three OFDM standards, i.e., LTE, DVB-H, and WiFi.

Hardware complexity was explored by comparing the proposed architecture to an equivalent accelerator. Even though the CGRA synchronization engine results in an overhead, its benefits in flexibility and scalability are evident. As the number of concurrent standards increases, this overhead becomes negligible, more so when the entire system is placed into perspective. Moreover, verification of the synchronization engine, the main element architectural element, was performed and measurements reported an average power dissipation of 1.9 mW when receiving WLAN standard and 1.6 mW during its configuration, both measured at 1.2 V.

16

Thesis summary

The design and implementation of mobile terminals is an extremely complex process, often involving a large number of disciplines, from mastering complicated mathematical models to proficiency of physical and electrical properties of specific materials involved in the manufacturing. In order to overcome this challenge, abstraction layers are used to separate the various disciplines, so that, a well trained individual can design this complex device from a system perspective. However, when the already complicated disciplines evolve, the abstraction layers become more and more complex. Thus, it becomes increasingly difficult for a single individual to make the right choices, resulting in sub-optimal implementations. In this thesis, three approaches to this interdisciplinary methodology were applied that resulted in the three parts of this thesis:

16.1. SIGN-BIT ESTIMATION

In the first approach, the co-design is executed in the direction of hardware architecture towards the algorithm. This is done by considering extreme quantization on the algorithm development. Quantization is more commonly considered a hardware aspect and is often neglected when designing the desired algorithm. However, a more realistic algorithm assessment can be obtained when quantization is considered in the initial design. In the approach taken in this thesis, it was shown that by carefully adapting the algorithm specification, the hardware complexity is dramatically reduced.

   More concisely, in Part I it is demonstrated that it is possible to estimate symbol timing, CFO, and SNR only with the use of a single bit or Sign-Bit of the received signal. The advantages and disadvantages in terms of hardware complexity and performance are shown. The capability to support several OFDM standards is demonstrated by the realization of a multi-standard ar-

chitecture. Performance simulations under realistic LTE scenarios indicate little or no degradation for symbol-timing and CFO estimation in comparison to higher resolution equivalent, whereas for SNR estimation the performance of SB is similar to its higher precision equivalent with much lower complexity. The potential of SB estimation techniques for low power receiver terminals was demonstrated with post-layout power simulations with back-annotated information.

To this effect, it is shown that SB-estimation architecture dissipated at least 1/70 of the power consumption compared to similar existing works in the open technical literature. Moreover, measurements of the time-frequency correlator are presented and show an average power consumption 0.62 mW for LTE with 20 MHz bandwidth and power supply of 1.2 V. To the best of the author's knowledge this is a new technique on algorithm development. Even though it cannot be applied to just any arbitrary application, in this work, it is demonstrated that it is possible to use in at least more than a single application.

16.2. CUSTOM DSP PROCESSING FOR LTE-A

With the introduction of CA in LTE-A, the maximum transmission bandwidth is increased to 100 MHz. This large bandwidth is not distributed uniformly among the carriers involved, but rather, there are large numbers of combinations that need to be supported. A large number of filtering modules is needed if all combinations are supported; one per combination. Thus, in the second approach of algorithm-architecture co-design, the optimization is done from the hardware side towards the algorithm side. For this, an extremely flexible and generic architecture, designed for SDR applications, is used to implement a selection of filtering algorithms. These filters are implemented as software instructions on the SDR platform. Once they are implemented on the SDR platform, the specialized modules used for filtering can be removed and the filtering can be done with DSP instructions instead.

In Part II on page 87 the analysis concludes that an SDR platform is a compelling alternative for the challenge imposed by LTE-A. The platform can dynamically change the filtering method in accordance with the required bandwidth. Simulations show that the number of data to be processed can be reduced up to 70% in the CCs that contain control signal. Possibly in future deployment of a pure LTE-A with no backward compatibility savings can be even higher. Moreover, the bandwidth reduction also contributes to sample rate reduction that is highly beneficial in the entire baseband. The maximum reduction accounts for 30 K samples for FFT filtering when reducing from 100 MHz to 1.080 MHz equivalent to the minimum assigned bandwidth of 6 RBs.

In addition, simulations show that the performance loss is marginal in terms of BER performance. Then, it is reasonable to conclude that the performance loss is relatively small and does not add significant noise to the posterior stages in the receiver. On the architectural side, it is important to note that the largest benefit for SDR platforms in terms of energy savings come from architecture optimization. Since FFT was the most optimized method, the number of required clock cycles dramatically decreases by using an FFT filtering method instead of a convolution method.

16.3. MULTI-STANDARD SUPPORT

In the third and last approach presented in Part III, the co-design is performed in the mid-point between algorithm and architecture. With the ultimate goal of supporting concurrent reception to more than one RAT, a special CGRA architecture is proposed so that algorithm efficiency can be dynamically traded-off with power consumption.

Algorithm-Architecture Co-Design plays a critical role in the design of future terminals since the level of flexibility required by the mobiles of the future, typical hardware accelerators, DSPs or general purpose processor are not sufficient as standalone solutions. When a mobile terminal is subject to support multiple-streams concurrently, the critical points becomes an architecture that has good flexibility with limited overhead.

Hardware complexity was explored by comparing the proposed architecture to an equivalent accelerator. Even though the CGRA synchronization engine results in an overhead, its benefits in flexibility and scalability are evident; as the number of concurrent standards increases, this overhead becomes negligible, more so when the entire system is placed into perspective. Moreover, verification of the synchronization engine, the main element architectural element, was performed and measurements reported an average power dissipation of 1.9 mW when receiving WLAN standard and 1.6 mW during its configuration, both measured at 1.2 V.

Appendix

Popular science summary

REMOVING BOUNDARIES BETWEEN THEORY AND PRACTICE, IN ORDER TO CONSTRUCT MORE EFFICIENT TELEPHONES. From the pure technological perspective (leaving the business aside), the design of electronic devices, such as mobile phones, is an extremely complex process often involving a large number of disciplines, from mastering complicated mathematical models to proficiency of physical and electrical properties of specific materials involved in the manufacturing. In order to overcome this challenge, abstraction layers are used to separate the various disciplines, so that, a well trained individual can design this complex device from a system perspective. However, as one can expect, when the already complicated disciplines evolve the abstraction layers become more and more complex, thus, it becomes increasingly difficult for a single individual to make the right choices, resulting in sub-optimal implementations.

In the thesis "Algorithm-Architecture Co-Design for Digital Front-Ends in Mobile Receivers", the author tries to defeat this challenge by removing the boundary between algorithm and hardware architecture. An algorithm, in this case, describes a mathematical model characterizing the procedure for all the mathematical operations that the mobile phone (or parts of it) need to perform in order to arrive to the desired functionality. The hardware architecture describes the phone's physical system, its components and the relation between those components.

In this thesis, the cross-disciplinary optimization is done in three manners: In *the first*, the optimization is done from architecture side toward the algorithm side. Here, the effects of extreme quantization, an effect typically seen as a hardware-only issue, is modeled in the algorithm development so that with just the right algorithm tinkering, parts of the mobile phone can save up to 80% in energy consumption. In *the second* manner, the optimization is done in the opposite direction, from the algorithm side towards the

hardware side. Here, the algorithm is set fixed, and a number of hardware architectures are selected for filtering. The correct selection leads to one of the hardware modules to become redundant. Then, by removing this redundant module, the system utilization is increased and the entire architecture becomes more flexible. In *the third* and last manner, the optimization is done just in the boundary of these disciplines. In order to support various standards in a mobile phone, the hardware architecture has to be as flexible as possible, in order to do so, a completely reconfigurable module was placed in the system that can run virtually any algorithm, by doing so, there a benefit in implementing in the same architecture: complex algorithm that consume high power, but perform highly advance operations, or simple algorithms that consume little power, but can only be used on quasi-ideal scenarios,

by Isael Diaz.

Acronyms

3GPP	3rd Generation Partnership Project	**BS**	Base Station
ADC	Analog to Digital Converter	**CA**	Carrier Aggregation
		CCA	Carrier Component Aggregation
ADRES	Architecture for Dynamically Reconfigurable Embedded Systems	**CC**	Carrier Component
		CFO	Carrier Frequency Offset
AFE	Analog Front-End	**CGA**	Coarse-Grain Architecture
AGC	Automatic Gain Control	**CGRA**	Coarse Grain Reconfigurable Architecture
AGRAC	Automatic Gain and Resource Activity Controller		
		CIC	Cascaded Integrator-Comb
ASIC	Application-Specific Integrated Circuit	**CLB**	Configurable Logic Block
ASIP	Application-specific instruction-set processor	**CORDIC**	COordinate Rotation DIgital Computer
AWGN	Additive White Gaussian Noise	**CP**	Cyclic Prefix
		DFE-Rx	Digital Front-End Receiver
BER	Bit Error Rate		

DFE-Tx	Digital Front-End Transmitter		**FHSS**	Frequency-Hopping Spread Spectrum
DFE	Digital Front-End		**FIFO**	First Input First Output
DFT	Discrete Fourier Transform		**FIR**	Finite Impulse Response
DL	Downlink		**FP7**	Seventh Framework Programme for Research
DMA	Direct Memory Access		**FPGA**	Field-Programmable Gate Array
DSP	Digital Signal Processor		**FSK**	Frequency Shift Keying
DSSS	Direct-Sequence Spread Spectrum		**FSM**	Finite State Machine
DVB-H	Digital Video Broadcasting - Handheld		**FU**	Functional Unit
			GI	Guard Interval
DVB-T	Digital Video Broadcasting - Terrestrial		**GPP**	General Purpose Processor
DVB	Digital Video Broadcasting		**GPS**	Global Positioning System
EPA	Extended Pedestrian A		**GSM**	Global System for Mobile communication
ETU	Extended Typical Urban		**HIPERLAN**	HIgh PERformance Local Area Network
EU	European Union		**HSPA**	High Speed Packet Access
EVA	Extended Vehicular A			
FCC	Federal Communications Commission		**ICI**	Inter-Carrier-Interference
			IC	Integrated Circuit
FDD	Frequency Division Duplexing		**IEEE**	Institute of Electrical and Electronics Engineers
FEC	Forward Error Correction			
FE	Front-End		**IFFT**	Inverse Fast Fourier Transform
FFT	Fast Fourier Transform		**IIR**	Infinite Impulse Response

ILP	Instruction Level Parallelism		**MISO**	Multiple Input Single Output
IMEC	Inter-university Microelectronics Research Center		**ML**	Maximum Likelihood
			MPSOC	Multi-Processor System on Chip
IP	Internet Protocol		**NFC**	Near-Field Communication
ISI	Inter-Symbol-Interference			
ITU	International Telecommunication Union		**NoC**	Network on Chip
			OFDM	Orthogonal Frequency Division Multiplexing
IoT	Internet of Things		**OSI**	Open Systems Interconnection
IrDA	Infrared Data Association			
LAN	Local Area Networks		**P/S**	Parallel to Serial
LDPC	Low-Density Parity-Check Code		**PAM**	Pulse Amplitude Modulation
			PDF	Probability Density Function
LOS	Line of Sigth			
LTE-A	Long Term Evolution - Advanced		**PDP**	Power Delay Profile
			PHY	Physical Layer
LTE	Long Term Evolution		**PLL**	Phase-Locked Loop
LTF	Long Training Field		**PRB**	Physical Resource Block
LTI	Linear Time Invariant		**PRE**	Physical Resource Element
LUT	Look-Up Table			
MAC	Medium Access Control		**PSD**	Power Spectral Density
MBSFN	Multimedia Broadcast Single Frequency Network		**PSK**	Phase Shift Keying
			PSS	Primary Synchronization Signal
MIMD	Multiple Instruction Multiple Data		**PSS**	Primary Synchronization Signal
MIMO	Multiple Input Multiple Output		**QAM**	Quadrature Amplitude Modulation

QAM	Quadrature Amplitude Modulation		**STF**	Short Training Field
QPSK	Quadrature Phase Shift Keying		**SoC**	System On Chip
			TDD	Time-Division Duplexing
QoS	Quality of Service			
RAM	Random Access Memory		**TPS**	Transmitter Parameter Signalling
RAM	Random Access Memory		**Tx**	Transmitter
RAT	Radio Access Technology		**UL**	Uplink
RB	Resource Block		**UMTS**	Universal Mobile Telecommunications System
RISC	Reduced Instruction Set Computing			
Rx	Receiver		**VHT**	Very High Throughput
S/P	Serial to Parallel		**VLIW**	Very Long Instruction Word
SB	Sign-Bit		**WCDMA**	Wide-band Code Division Multiple Access
SDR	Software Defined Radio			
SFN	Single Frequency Network		**WLAN**	Wireless Local Area Network
SFO	Sampling Frequency Offset		**WPAN**	Wireless Personal Area Network
SIMD	Single Instruction Multiple Data		**WP**	Work Package
SIMO	Single Input Multiple Output		**WWAN**	Wireless Wide Area Network
SIR	Signal to Interference Ratio		**WiFi**	IEEE 802.11
SISO	Single Input Single Output		**WiMAX**	Worldwide Interoperability for Microwave Access
SNR	Signal-to-Noise Ratio		**c.c.**	clock cycle
SSS	Secondary Synchronization Signal		**i.i.d.**	independent and identically distributed
			μ**C**	Micro-Controller

Glossary of symbols

$*$	Convolution operator.
$\|\{\}\|$	Absolute value.
α	QAM modulated data sub-carrier.
β	Output of complex moving sum in auto-correlation function.
$\lceil\{\}\rceil$	Ceiling function.
$\Delta\varphi$	Carrier frequency offset normalized to the sub-carrier spacing.
δ	Dirac delta function.
Δ_f	Sub-carrier spacing.
η	Noise power.
Γ	Number of contiguous resource blocks assigned to a particular user in LTE.
γ	Output of auto-correlation function.
$\hat{\{\}}$	Estimated value.
$\overline{\{\}}$	Complex conjugation operator.
σ_s^2	Variance of signal s.
σ_w^2	Variance of the complex signal w.
θ	Symbol start.

Υ Output of the complex subtraction in SNR estimation.

ζ Symbol end.

$\{\}^q$ Quantized signal

A Number of elements to be added in a moving sum.

B Transmission bandwidth.

$E[]$ Expectation operation.

F Modulation function based on IFFT.

f_c Carrier Frequency

F_s Sample-rate.

G Transfer function in the z-domain.

I Power of inter-carrier-interference.

L Length of CP in number of samples.

M Differential delays in CIC filters.

N FFT size and number of samples in OFDM symbol.

p probability that the sign-bit subtraction yields a result different from zero.

$Q()$ Q function, also known as tail probability function.

R Decimation factor.

r Complex received signal at baseband.

S Signal power.

s Complex transmitted signal at baseband.

S_D Doppler spectrum.

SNR Signal-to-noise ratio.

T_s Duration of OFDM symbol including CP.

T_u Duration of OFDM symbol without CP.

T_{CP} Duration of cyclic prefix.

ν_D Doppler shift in Hertz.

w Complex Gaussian noise

Bibliography

[1] "The world in 2013, facts and figures." ITU/ICT, February 2013.

[2] Ericsson, "White paper: More than 50 billion connected devices," white paper, Ericsson AB, February 2011.

[3] J. M. Rabaey, "The swarm at the edge of the cloud-a new perspective on wireless," in *VLSI Circuits (VLSIC), 2011 Symposium on*, pp. 6–8, IEEE, 2011.

[4] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[5] G. Moore, "Cramming more components onto integrated circuits," *Electronics Magazine*, vol. 38, April 1965.

[6] S. Koenig, D. Lopez-Diaz, J. Antes, F. Boes, R. Henneberger, A. Leuther, A. Tessmann, R. Schmogrow, D. Hillerkuss, R. Palmer, *et al.*, "Wireless sub-thz communication system with high data rate," *Nature Photonics*, vol. 7, no. 12, pp. 977–981, 2013.

[7] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Wiley series in telecommunications, Cambridge University Press, 2005.

[8] J. Proakis and M. Salehi, *Digital Communications*. McGraw-Hill International Edition, McGraw-Hill Higher Education, 2008.

[9] A. Molisch, *Wireless Communications*. Wiley, 2005.

[10] O. Edfors, M. Sandell, J.-J. van de Beek, D. Landström, and F. Sjöberg, "An introduction to orthogonal frequency-division multiplexing," Tech. Rep. Research report/1996:16, Luleå University of Technology, Sept. 1996.

[11] Y. Li, *Orthogonal Frequency Division Multiplexing for Wireless Communications*. Springer, 2006.

[12] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS Long Term Evolution: From Theory to Practice*. John Wiley and Sons, 2009.

[13] O. Edfors, *Low-Complexity Algorithms in Digital Receivers*. PhD thesis, Luleå University of Technology, Sep. 1996.

[14] J. Löfgren and P. Nilsson, "On hardware implementation of radix 3 and radix 5 FFT kernels for LTE systems," in *NORCHIP*, pp. 1–4, IEEE, 2011.

[15] E. Dubois and A. Venetsanopoulos, "A new algorithm for the radix-3 FFT," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 26, no. 3, pp. 222–225, 1978.

[16] M. Speth, S. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for OFDM-based broadband transmission part II. A case study," *IEEE Transactions on Communications*, vol. 49, pp. 571–578, April 2001.

[17] P. H. Moose, "A technique for orthogonal frequency division multiplexing frequency offset correction," *IEEE Transactions on Communications*, vol. 42, pp. 2908–2914, October 1994.

[18] F. Classen and H. Meyr, "Frequency synchronization algorithms for OFDM systems suitable for communication over frequency selective fading channels," in *IEEE 44th Vehicular Technology Conference*, pp. 1655–1659, IEEE, 1994.

[19] H. Nogami and T. NAGASHIMA, "A frequency and timing period acquisition technique for OFDM systems," *IEICE Transactions on Communications*, vol. 79, no. 8, pp. 1135–1146, 1996.

[20] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for OFDM," *IEEE Transactions on Communications*, vol. 45, pp. 1613–1621, December 1997.

[21] J. van de Beek, M. Sandell, and P. Börjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1800–1805, July 1997.

[22] E. Amos, "Mobile communication system," May 16 1972. US Patent 3,663,762.

[23] W. Lee, J.-S. Kwak, P. Kafle, J. Tingleff, T. Yucek, R. Porat, V. Erceg, Z. Lan, and H. Harada, "IEEE P802.11 Wireless LANs TGaf PHY proposal," tech. rep., IEEE LAN/MAN Standards Committee, 2012.

[24] IMEC, "Next generation air interfaces," tech. rep., IMEC, 2007.

[25] E. Dahlman, S. Parkvall, and J. Sköld, *LTE/LTE-Advanced for Mobile Broadband.* Academic Press, 2011.

[26] 3GPP, "TS 36.211 V8.3.0 Physical Channels and Modulation," tech. rep., 3rd Generation Partnership Project, May 2008.

[27] 3GPP, "TR36.815 V9.1.0 LTE-Advanced feasibility studies in RAN WG4," tech. rep., 3rd Generation Partnership Project, Jun. 2010.

[28] IEEE, "802.11ac std., part 11: Wireless LAN medium access control and physical layer specifications," tech. rep., IEEE LAN/MAN Standards Committee, 2013.

[29] IEEE, "IEEE P802.11N/D2.00," tech. rep., IEEE LAN/MAN Standards Committee, Feb. 2007.

[30] ETSI, "EN 302 304v1.1.1: Transmission system for handheld terminals," tech. rep., ETSI, 2004.

[31] ETSI, "ETSI EN 300 744: Digital Video Broadcasting (DVB); Framing structure, channel coding and modulation for digital terrestrial television (DVB-T)," tech. rep., ETSI, 2004.

[32] ETSI, "TR 102 377 v1.2.1: DVB-H implementation guidelines," tech. rep., ETSI, 2005.

[33] G. Constantinides and G. Woeginger, "The complexity of multiple wordlength assignment," *Applied Mathematics Letters*, vol. 15, no. 2, pp. 137 – 140, 2002.

[34] G. A. Constantinides, P. Y. K. Cheung, and W. Luk, "Multiple precision for resource minimization," in *Proceedings of the 2000 IEEE Symposium on Field-Programmable Custom Computing Machines*, FCCM '00, (Washington, DC, USA), pp. 307–, IEEE Computer Society, 2000.

[35] L. Zhang, Y. Zhang, and W. Zhou, "Floating-point to Fixed-point Transformation Using Extreme Value Theory," in *Eighth IEEE/ACIS International Conference on Computer and Information Science (ICIS)*, pp. 271–276, June 2009.

[36] J. Zhang, Z. Zhang, S. Zhou, M. Tan, X. Liu, X. Cheng, and J. Cong, "Bit-level optimization for high-level synthesis and FPGA-based acceleration," in *Proceedings of the 18th annual ACM/SIGDA International Symposium on Field Programmable Gate Arrays*, pp. 59–68, February 2010.

[37] F. Horlin and A. Bourdoux, *Digital compensation for analog front-ends: a new approach to wireless transceiver design.* Wiley, 2008.

[38] S. Attallah, Y. Wu, and J. Bergmans, "Low complexity blind estimation of residual carrier offset in orthogonal frequency division multiplexing based," *IET Communications*, vol. 1, pp. 604–611, August 2007.

[39] S. LMAI, A. Bourre, C. Laot, and S. Houcke, "An efficient blind estimation of carrier frequency offset in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. PP, pp. 1–1, October 2013.

[40] M.-F. Sun, J.-Y. Yu, and T.-Y. Hsu, "Estimation of Carrier Frequency Offset With I/Q Mismatch Using Pseudo-Offset Injection in OFDM Systems," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 55, pp. 943–952, April 2008.

[41] F. Classen and H. Meyr, "Frequency synchronization algorithms for OFDM systems suitable for communication over frequency selective fading channels," in *IEEE 44th Vehicular Technology Conference (VTC)*, vol. 3, pp. 1655–1659, June 1994.

[42] W. Fan and C.-S. Choy, "Robust, Low-Complexity, and Energy Efficient Downlink Baseband Receiver Design for MB-OFDM UWB System," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 59, pp. 399–408, February 2012.

[43] 3GPP, "TS 36.213 v8.3.0 Physical Layer Procedures," tech. rep., 3rd Generation Partnership Project, May 2008.

[44] 3GPP, "R4-070572 Proposal for LTE Channel Models," tech. rep., 3rd Generation Partnership Project Std, May 2007.

[45] T. Keller and L. Hanzo, "Orthogonal frequency division multiplex synchronisation techniques for wireless local area networks," in *Seventh IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, vol. 3, pp. 963–967, October 1996.

[46] L. Wilhelmsson, I. Diaz, T. Olsson, and V. Öwall, "Performance Analysis of Sign-Based Pre-FFT Synchronization in OFDM Systems," in *IEEE 71st Vehicular Technology Conference - Spring (VTC)*, pp. 1–5, May 2010.

[47] J. H. Stott, "The effects of frequency errors in OFDM," tech. rep., BBC RD, 1995.

[48] F. Tufvesson, O. Edfors, and M. Faulkner, "Time and frequency synchronization for OFDM using PN-sequence preambles," in *IEEE VTS 50th Vehicular Technology Conference - Fall (VTC)*, vol. 4, pp. 2203–2207, September 1999.

[49] A. D. Whalen, *Detection of signals in noise*. New York : Academic Press, 1971.

[50] N. C. Beaulieu, "A useful integral for wireless communication theory and its application to rectangular signaling constellation error rates," *IEEE transactions on communications*, vol. 54, no. 5, pp. 802–805, 2006.

[51] A. Prudnikov, Y. Brychkov, and O. Marichev, *Integrals and series*. Gordon and Breach Science Publishers, 1992.

[52] L. Vachhani, K. Sridharan, and P. Meher, "Efficient CORDIC Algorithms and Architectures for Low Area and High Throughput Implementation," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 56, pp. 61–65, January 2009.

[53] P. Meher, J. Valls, T.-B. Juang, K. Sridharan, and K. Maharatna, "50 Years of CORDIC: Algorithms, Architectures, and Applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 56, pp. 1893–1907, September 2009.

[54] I. Diaz, L. Wilhelmsson, J. Rodrigues, J. Löfgren, T. Olsson, and V. Öwall, "A sign-bit auto-correlation architecture for fractional frequency offset estimation in OFDM," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, June 2010.

[55] J. Cho, Y. Cho, M. Islam, J. Kim, and W.-K. Cho, "Hardware-efficient auto-correlation for synchronization of MIMO-OFDM WLAN systems," in *International SoC Design Conference (ISOCC)*, pp. 560–563, November 2009.

[56] A. Troya, K. Maharatna, M. Krstic, E. Grass, U. Jagdhold, and R. Kraemer, "Low-Power VLSI Implementation of the Inner Receiver for OFDM-Based WLAN Systems," *IEEE Transactions on Circuits and Systems I*, vol. 55, pp. 672–686, March 2008.

[57] J. Cho, J. Kim, and W. Cho, "VLSI Implementation of Auto-Correlation Architecture for Synchronization of MIMO-OFDM WLAN Systems," *Journal of Semiconductor Technology and Science*, vol. 10, September 2010.

[58] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, "LTE-advanced-evolving LTE towards IMT-advanced," in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, pp. 1–5, IEEE, 2008.

[59] F. K. Jondral, "Software-defined radio: basics and evolution to cognitive radio," *EURASIP Journal on Wireless Communications and Networking*, vol. 2005, no. 3, pp. 275–283, 2005.

[60] F. Agnelli, G. Albasini, I. Bietti, A. Gnudi, A. Lacaita, D. Manstretta, R. Rovatti, E. Sacchi, P. Savazzi, F. Svelto, *et al.*, "Wireless multi-standard terminals: system analysis and design of a reconfigurable RF front-end," *IEEE Circuits and Systems Magazine*, vol. 6, no. 1, pp. 38–59, 2006.

[61] F. Bouwens, M. Berekovic, A. Kanstein, and G. Gaydadjiev, "Architectural exploration of the ADRES coarse-grained reconfigurable array," in *Reconfigurable Computing: Architectures, Tools and Applications*, pp. 1–13, Springer, 2007.

[62] M. J. Smith, *Digital Filtering: A Computer Laboratory Textbook*. New York, NY, USA: John Wiley & Sons, Inc., 1993.

[63] M. P. Donadio, "CIC filter introduction." Available in `http://www.dspguru.net/sites/dspguru/files/cic.pdf`, 2000. Last time accessed : March 2, 2014.

[64] J. W. Cooley and J. W. Tukey, "An algorithm for the machine calculation of complex fourier series," *Math. comput*, vol. 19, no. 90, pp. 297–301, 1965.

[65] T. Lenart and V. Öwall, "Architectures for dynamic data scaling in 2/4/8K pipeline FFT cores," *Transactions on Very Large Scale Integration Systems*, vol. 14, no. 11, pp. 1286–1290, 2006.

[66] L. Van der Perre, B. Bougard, J. Craninckx, W. Dehaene, L. Hollevoet, M. Jayapala, P. Marchal, M. Miranda, P. Raghavan, T. Schuster, *et al.*, "Architectures and circuits for software-defined radios: Scaling and scalability for low cost and low energy," in *IEEE International Solid-State Circuits Conference, ISSCC. Digest of Technical Papers.*, pp. 568–569, IEEE, 2007.

[67] B. Bougard, B. De Sutter, D. Verkest, L. Van der Perre, and R. Lauwereins, "A Coarse-Grained Array Accelerator for Software-Defined Radio Baseband Processing," *IEEE Micro*, vol. 28, july-aug. 2008.

[68] B. Mei, S. Vernalde, D. Verkest, H. De Man, and R. Lauwereins, "ADRES: An architecture with tightly coupled VLIW processor and coarse-grained reconfigurable matrix," in *Field Programmable Logic and Application*, pp. 61–70, Springer, 2003.

[69] M. Iwamura, K. Etemad, M.-H. Fong, R. Nory, and R. Love, "Carrier aggregation framework in 3GPP LTE-advanced [WiMAX/LTE Update]," *Communications Magazine*, 2010.

[70] F. Bouwens, M. Berekovic, B. De Sutter, and G. Gaydadjiev, "Architecture enhancements for the ADRES coarse-grained reconfigurable array," in *HiPEAC*, 2008.

[71] M. Palkovic, P. Debacker, P. Avasare, S. Dupont, and T. V. Aa, "Power Estimation at Different Abstraction Levels for Wireless Baseband Processors.," *J. Low Power Electronics*, vol. 8, no. 5, pp. 726–738, 2012.

[72] A. Baschirotto, R. Castello, F. Campi, G. Cesura, M. Toma, R. Guerrieri, R. Lodi, L. Lavagno, and P. Malcovati, "Baseband analog front-end and digital back-end for reconfigurable multi-standard terminals," *IEEE Circuits and Systems Magazine*, vol. 6, pp. 8–28, Quarter 2006.

[73] K. van Berkel, F. Heinle, P. P. E. Meuwissen, K. Moerman, and M. Weiss, "Vector processing as an enabler for software-defined radio in handheld devices," *EURASIP Journal of Applied Signal Processing*, vol. 2005, pp. 2613–2625, 2005.

[74] Y. Lin, Y. L. Hyunseok, M. Woh, Y. Harel, S. Mahlke, T. Mudge, and C. Chakrabarti, "SODA: A Low-power Architecture For Software Radio," in *In Proc. of the 33rd Annual International Symposium on Computer Architecture*, pp. 89–101, 2006.

[75] "About Technikon." Available in `http://www.technikon.com/`, 2014. Last time accessed : January 15, 2014.

[76] "Ericsson AB, company facts." Available in `http://www.ericsson.com/thecompany/company_facts`, 2013. Last time accessed : January 15, 2014.

[77] "About Infineon Austria." Available in `http://www.infineon.com/cms/regional-pages/infineon-austria/en/about.html`, 2013. Last time accessed : January 15, 2014.

[78] "About Lund University." Available in `http://www.lunduniversity.lu.se/about-lund-university`, 2013. Last time accessed : January 15, 2014.

[79] "Linköping University, facts and figures." Available for download in `http://www.liu.se/om-liu/siffror/siffror3/1.456846/fickfakta_engelsk_webb_130417.pd`, April 2013. Last time accessed : January 15, 2014.

[80] "About IMEC." Available in `http://www2.imec.be/be_en/about-imec.html`, 2014. Last time accessed : January 15, 2014.

[81] "About KU Leuven, facts and figures." Available in `http://www.kuleuven.be/about/`, 2013. Last time accessed : January 15, 2014.

[82] O. Edfors, V. Öwall, I. Diaz, F. Foroughi, T. Olsson, L. Wilhelmsson, J. Svensson, L. v.d. Perre, D. Novo, M. Palkovic, and L. Hollevoet, "D2.2 architecture-aware algorithmic solution(s) including data type refinement," tech. rep., Multibase Consortium, September 2009.

[83] L. v.d. Perre, L. Hollevoet, L. Min, R. Fasthuber, M. Stala, T. Olsson, L. Wilhelmsson, I. Diaz, O. Edfors, V. Öwall, F. Foroughi, D. Liu, R. Asghar, D. Wu, and W. Dehaene, "D4.2 identification and architectural development of key building blocks," tech. rep., Multibase Consortium, January 2009.

[84] M. A. Mazidi, J. Mazidi, R. McKinlay, and P. Ingendorf, *PIC Microcontroller and Embedded Systems*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2005.

[85] C. W. Farrow, "A continuously variable digital delay element," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 2641–2645 vol.3, 1988.

[86] T. Lenart, *Design of Reconfigurable Hardware Architectures for Real-time Applications: Modeling and Implementation*. PhD thesis, Lund University, June 2008.

[87] C. Zhang, I. Diaz, P. Andersson, J. Rodrigues, and V. Öwall, "Reconfigurable cell array for concurrent support of multiple radio standards by flexible mapping," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1696–1699, 2011.