Diss. ETH No. 28852

Extended Coverage Synchronization and Positioning for the Cellular Internet of Things

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES of ETH ZURICH
(Dr. sc. ETH Zurich)

presented by

MAURO ANTONIO SALOMON
MSc ETH EEIT, ETH Zurich
born on 14.12.1990
citizen of Sierre VS, Switzerland

accepted on the recommendation of

Prof. Dr. Qiuting Huang, examiner
Prof. Dr. Andreas Burg, co-examiner
Prof. Dr. Matthias Korb, co-examiner

2022

Acknowledgments

This thesis would not have been possible without the support of many people in my professional and private circle. Inevitably, I will not mention everyone below, but they all deserve my sincere gratitude. I would first like to thank Prof. Qiuting Huang for the opportunity to work on such an exciting project and for his guidance throughout my years at the Integrated Systems Laboratory (IIS) of ETH Zurich. Special acknowledgments go to Prof. Andreas Burg from EPFL for reading and co-examining my thesis. I valued the opportunity to expose my work to his expertise in VLSI design and digital signal processing. I also express my deepest gratitude to Prof. Matthias Korb for co-supervising and examining the thesis. His extensive support and advice have been priceless during my whole time as a doctoral student. Furthermore, he was a key contributor to our research activities.

Among my academic colleagues, I would especially like to thank my fellow doctoral student and J90 office mate Stefan Lippuner for his feedback on the first draft of the thesis, his immense contributions to the modem system-on-chips, and our countless and fruitful discussions at the coffee corner. Special thanks also to Benjamin Weber, who supervised my Master's thesis, for the initial recommendation for the PhD position, and to Harald Kröll for his involvement in the genesis of the SymbIoT project. I also want to acknowledge the support of all the outstanding engineers of ACP AG that contributed to the development of the system-on-chips presented in this work: David Tschopp, Samuel Willi, Michael Gautschi, Stefan Altorfer, Thomas Dellsperger, Jürgen Rogin, Edwin Dornbierer, Andreas Traber, and the list is probably incomplete.

At this point, I would also like to thank the group of people who made the IIS a great place to work and do academic research during my years there. Thomas Burger for managing and fairly distributing the teaching workload. Frank Gürkaynak and Beat Muheim from the Microelectronics Design Center for their support with all the EDA tools involved in the IC design process. Christoph Wicki and Adam Feign for managing the IT infrastructure. Thomas Kleier and Hansjörg Gisler for the hands-on support, especially when soldering was involved. Monika Zwahlen for all the administrative matter that was completely transparent to me as an ETH employee. Among the faithful members of the 11:30 lunch group, I would like to mention Noé Brun, Philipp Schönle, Florian Glaser, Guillaume Mocquard and Jérémy Ghichemerre, and Pirmin Vogel for always being cheerful companions.

Finally, I am deeply grateful to my family in Switzerland and Argentina and friends from Valais and Zurich. I enjoyed each shared moment, conversation, and trip that allowed me to take a breath and renew my motivation. My grandparents, uncles, aunts, cousins, and future in-laws have been a valuable source of advice and inspiration throughout my studies. The unconditional support of my parents, Antonio and Nora, and my siblings, Alessio and Sabrina, were precious for all my achievements. Last but not least, the love and patience of my fiancée, Sylvie, are the pillars that carried me through my doctorate.

February 2023

Abstract

The fast growth of the Internet of things (IoT) has triggered the emergence of new needs in wireless communications. The requirements of applications that rely on autonomous embedded systems differ from those of human-centric services that have traditionally dominated data traffic. The urge for facilitated wireless connectivity for small, low-cost, and energy-constrained devices has led to the rapid development of low-power wide area networks (LPWAN) over recent years. High connection density, low device complexity, low-power operating modes, ubiquitous coverage, and positioning capabilities are all essential features that LPWAN must support to satisfy the needs of machine-type communications (MTC).

Legacy cellular networks offer a global infrastructure that can fulfil the needs of all kinds of IoT use cases if it evolves to include a specific set of features for MTC. The standardization of Narrowband Internet-of-Things (NB-IoT) and enhanced Machine-Type Communication (eMTC) based on fourth-generation (4G) mobile technologies has provided cellular operators with sufficient tools to upgrade their networks for present and future machine-centric applications. However, the success of cellular IoT communications relies on the availability of low-cost modems that can take full advantage of the features included in the mentioned standards.

This thesis identifies the main challenges of developing such devices and provides solutions that fit the constraints of integrated systems. Amid the desired features, coverage extension is essential to achieve global availability without increasing the density of the deployed base stations. Specially tailored algorithms that deal with the reduced

capabilities of IoT equipment while fully exploiting the allocated radio resources are crucial to expand the reach of communication and positioning services. Repeated transmissions can compensate for the additional path loss suffered in remote or underground locations and increase signal coverage at the cost of higher latency and reduced throughput, both of which are acceptable for MTC. However, accurate synchronization is needed to minimize accumulation losses and decrease the overhead of repetitions.

Synchronization to the network needs the precise evaluation of errors in the local time and frequency references of devices to periodically correct residual offsets. The estimation strategies developed in this work and implemented on highly-integrated modem system-on-chips (SoCs) allow the communication with eMTC at input power levels as low as $-136.8\,\mathrm{dBm}$, the equivalent of $-24.8\,\mathrm{dB}$ signal-to-noise ratio (SNR). Furthermore, cellular networks supporting positioning based on the observed time difference of arrival (OTDOA) provide an alternative to the Global Navigation Satellite System (GNSS) and extend the coverage of location services to indoor and deep urban scenarios. The prototype modem SoC developed in this work showed an OTDOA positioning accuracy below $70\,\mathrm{m}$ in a laboratory setup using NB-IoT signals with cells at $-15\,\mathrm{dB}$ SNR.

Résumé

La croissance rapide de l'Internet des objets (IdO) a mis en lumière de nouvelles exigences en matière de communication. Les applications basées sur des systèmes embarqués autonomes n'ont pas les mêmes besoins que celles centrées sur les êtres humains qui ont traditionnellement dominé le traffic de données. La nécessité de simplifier l'accès à la connectivité sans fil pour des appareils de petite taille, à bas coût et à l'utilisation d'énergie réduite est à l'origine du développement récent de résaux étendus à basse consommation (LPWAN). Une haute densité de connections, des terminaux simplifiés, des modes de consommation réduite, une couverture globale et la géolocalisation sont autant d'éléments essentiels qui doivent être inclus dans les LPWAN pour satisfaire les besoins de la communication intermachines.

Les réseaux mobiles conventionnels offrent une structure mondiale qui peut remplir les exigences de toutes sortes de cas de figure dans l'IdO, pour autant que certaines améliorations y soient apportées. Avec l'introduction de Narrowband Internet-of-Things (NB-IoT) et enhanced Machine-Type Communication (eMTC), tous deux basés sur la technologie mobile de quatrième génération (4G), les dernières normes de communication offrent aux opérateurs de réseaux suffisamment d'outils pour adapter leur infrastructure aux besoins actuels et futurs des applications centrées sur les machines. Cependant, le succès des communications cellulaires pour l'IdO repose sur la disponibilité de modems à faible coût qui peuvent tirer pleinement parti des fonctionnalités permises par les standards nommés précédemment.

Cette thèse identifie les principaux défis de l'implémentation de tels dispositifs et fournit des solutions adaptées aux contraintes de systèmes intégrés. Parmi les caractéristiques souhaitées, l'extension

de la couverture du signal radio est essentielle pour atteindre une disponibilité globale sans augmenter la densité des stations de base déployées. Afin d'étendre la portée des services de communication et de localisation, il est crucial de concevoir des algorithmes qui exploitent pleinement les signaux reçus malgré les capacités réduites des équipements pour l'IdO. Des transmissions répétées peuvent compenser l'affaiblissement supplémentaire subi par les signaux sans fil pour atteindre des endroits éloignés ou souterrains. La couverture peut ainsi être augmentée au prix d'une latence plus élevée et d'un débit réduit, tous deux acceptables dans de nombreux cas de figure de communication intermachines. Cependant, une synchronisation précise est indispensable pour minimiser les pertes lors de l'accumulation de signaux et ainsi optimiser l'utilisation des répétitions.

La synchronisation au réseau nécessite l'évaluation précise de décalages de temps et de fréquence dans les dispositifs pour corriger périodiquement les erreurs résiduelles de leurs références. Les stratégies d'estimation développées dans ce travail et implémentées dans des modems intégrès en tant que systèmes sur puce (SoC) permettent de communiquer à l'aide d'eMTC à des niveaux de signal reçu aussi bas que $-136.8\,\mathrm{dBm}$, soit l'équivalent d'un rapport signal sur bruit (SNR) de $-24.8\,\mathrm{dB}$. En outre, les réseaux cellulaires qui supportent le positionnement basé sur la différence de temps d'arrivée observée (OTDOA) constituent une alternative au système mondial de navigation par satellite (GNSS) et étendent la couverture des services de localisation à des appareils déployés à l'intérieur de bâtiments ou dans des zones hautement urbanisées. Des expériences en laboratoire ont également démontré qu'une précision inférieure à $70\,\mathrm{m}$ peut être atteinte en utilisant le positionnement OTDOA avec NB-IoT, malgré des signaux de cellules à $-15\,\mathrm{dB}$ SNR.

Contents

Chapter 1

Introduction

Cellular networks are an essential component of the social and technological revolutions triggered by the advent of mobile phones. The importance of wireless connectivity has led to the worldwide deployment of cellular systems, bringing the Internet into everyone's hands. Over the last decade, the broad cellular coverage has contributed to the growth of the Internet of things (IoT) [1], an expanding network of objects connected to the Internet. The increased number of device-to-device links has highlighted the different requirements of machine-type communications (MTC) compared to human-centric scenarios and has driven the emergence of competing wireless communication standards.

With their global presence, cellular networks already provide a viable solution for IoT connectivity in many cases. However, past technologies target mainly personal use and miss the flexibility to accommodate all emerging applications. The evolution of cellular standards, from 2G to 4G, focused principally on improving user experience, e.g. by increasing data rates and reducing latencies. Nevertheless, these features are not always relevant for MTC. For example, remote sensing and monitoring only need limited throughput and can deal with long delays in data transfer. Instead, they would benefit from increased battery life to reduce operational costs and extended signal coverage to reach underground locations. 5G networks were developed to fulfil such requirements but need performant devices that fully exploit their features.

1.1 5G Standards for the IoT

The International Telecommunication Union (ITU), an agency of the United Nations that regulates communication and information technologies, decidedto group a set of protocols satisfying the needs of future wireless communications into a family of standards called the International Mobile Telecommunications-2020 (IMT-2020). The ITU fixed the requirements to be part of the IMT-2020 specifications in 2017 [2]. The three classes of use cases depicted in Fig. 1.1 were defined to cover the broad spectrum of target applications and to deal with their conflicting requirements [3]:

- enhanced mobile broadband (eMBB),
- ultra-reliable and low-latency communication (URLLC),
- massive machine-type communication (mMTC).

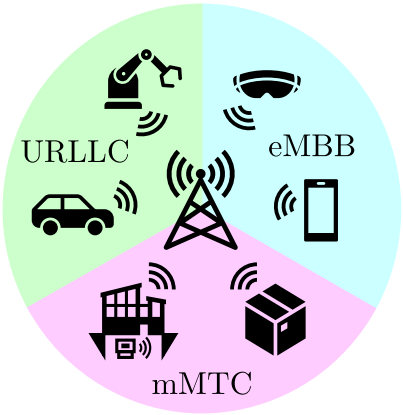

Figure 1.1: Classes of use cases considered for 5G cellular technologies as defined in the requirements for IMT-2020 [2].

The first scenario, eMBB, corresponds to the evolution of legacy mobile broadband to support even higher data rates. Its characteristics can also cover the needs of emerging applications, like virtual and

augmented reality, that would benefit from higher throughputs than supported by previous generations of mobile communication standards. The second, URLLC, was meant for so-called mission-critical services with stringent requirements on reliability and latency. For instance, traffic safety for autonomous cars and factory automation could rely on this kind of communication. Finally, mMTC gathers use cases where wireless connectivity should be accessible for numerous low-cost devices that only need small data rates, like packet or asset trackers, smart meters, and sensors for remote monitoring.

The 3rd Generation Partnership Project (3GPP) is the principal organization developing specifications for mobile communications and was already responsible for the third and fourth generations of cellular technologies. To comply with the objectives set by ITU, the 3GPP proposed a set of radio access technologies (RATs) that form its fifth-generation (5G) family of standards and were included in the first IMT-2020 release in 2021 [4]. In 3GPP's 5G portfolio, the eMBB and URLLC scenarios are covered by a newly introduced RAT named New Radio (NR) [3], while mMTC support is provided by two solutions, namely Narrowband Internet-of-Things (NB-IoT) and enhanced Machine-Type Communication (eMTC)[1] [5], both based on Long Term Evolution (LTE), the fourth-generation (4G) cellular standard.

Besides remote sensing and control, many IoT applications rely on tracking the geographical position of objects. In this case, knowing the target device's location is essential. Cellular networks facilitate position estimations in different ways. For example, NB-IoT and eMTC can serve as support for the Global Navigation Satellite System (GNSS) with the transmission of assistance information to the end device [6]. Furthermore, cellular signals received from different base stations (BSs) allow the estimation of the location of modems by multilateration. This last option uses the observed time difference of arrival (OTDOA) technique and is already part of the LTE specifications [7]. For IoT devices, OTDOA support was also included in NB-IoT and eMTC with improvements to extend its coverage.

[1] eMTC is also called LTE-M or LTE category M (Cat-M).

1.1.1 Challenges of Massive MTC

NB-IoT and eMTC have similar attributes and were designed to fulfil the main requirements of mMTC. The main features that make them suitable for such scenarios are:

- extending signal coverage,
- allowing low-power operation,
- enabling low-complexity devices, and
- supporting positioning services.

Notable differences between the two are an even lower complexity for NB-IoT and higher data rates and better support for mobility for eMTC. A new generation of cellular modems was needed to fully exploit the specific characteristics of these standards and unlock the potential of mMTC. Their development must focus on achieving a high integration level to minimize cost, efficiently managing power consumption to extend battery life, and ensuring ubiquitous connectivity and positioning.

1.1.2 Adoption of NB-IoT and eMTC

In 2016, 3GPP introduced the first version of NB-IoT and eMTC in Release 13 of LTE specifications. Since then, newer releases have updated and expanded both technologies. By now, the two are commercially available in numerous networks around the world, as depicted in Fig. 1.2. Their inclusion in the IMT-2020 standards and their compatibility with 4G and 5G networks are significant advantages compared to older cellular technologies used for IoT applications in the past. Furthermore, 2G and 3G support is being discontinued in an increasing number of regions [8, 9].

However, the adoption of NB-IoT and eMTC was not immediate and Fig. 1.2 also shows that many territories are not covered by either of the two RATs, yet. Many countries in Latin America, Asia, and Africa still rely on 2G for low-cost devices, without guarantees of coming deployments. Furthermore, the new cellular IoT (cIoT) standards had a fragmented adoption across countries, with some

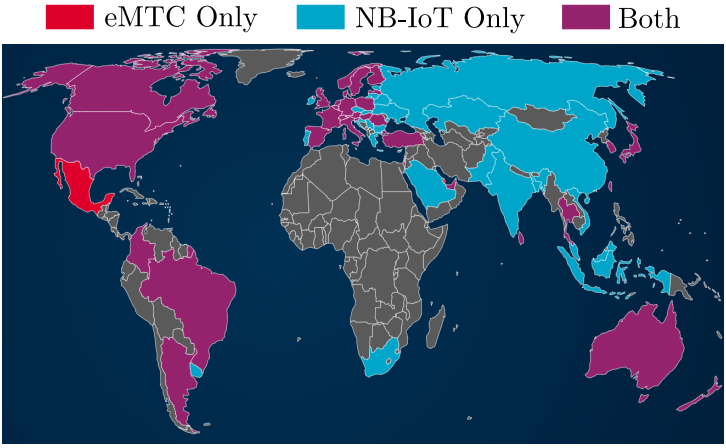

Figure 1.2: The GSMA's map of NB-IoT and eMTC commercial deployments as of February 2022 [10] © GSM Association 1999 - 2019.

networks supporting only one of the two solutions and providing irregular regional coverage. Thus, true global connectivity can only be achieved by multi-mode devices supporting NB-IoT, eMTC, and 2G for fallback. The need to support several technologies combined with high pressure on module size and cost is another constraint that has to be considered for appropriate system design.

1.2 Previous and Related Work

In past years, extensive research toward fully integrated cellular modems has been undertaken at the Integrated Systems Laboratory (IIS) of ETH Zurich. Collaboration with the project partner ACP AG enabled the implementation of a solution including radio frequency (RF) subsystem and digital baseband (DBB) processing for 2G into a single integrated circuit [11]. A step further was taken in [12] by also including a microprocessor system into a combined 2G and NB-IoT modem, integrated into a single dual-mode RF-system-on-chip (SoC) caple of hosting the cellular protocol stacks and a user application.

This work builds on the mentioned research and is the result of a close collaboration with the industrial partner and fellow doctoral students and researchers from the IIS. Among them, Stefan Lippuner widely contributed to the development of the triple-mode modem SoC presented in this thesis [13]. Besides his work on the architecture and back-end implementation of the SoC, he was responsible for crucial tasks in the digital baseband that are complementary to the material presented in this work. Especially, he designed algorithms that drastically improve cell search, initial synchronization, and data decoding in extended coverage [14].

1.3 Contributions

Among the desired features of mMTC, coverage enhancement has a significant impact on signal processing and imposes considerable challenges on the physical layer (PHY) implementation of modems [15]. In extended coverage conditions, when the power reaching the device is well below the thermal noise floor, establishing and maintaining a communication link requires special approaches that are not needed for legacy mobile broadband. This work proposes solutions to develop a multi-mode cIoT modem with extended coverage and positioning support. The presented algorithms were integrated into a single SoC while maintaining device complexity low. The performance gains achieved with improved signal processing were further exploited to reduce power consumption.

Time and frequency synchronization between the receiver and the transmitter is essential in wireless positioning and communication systems. The reception of cellular signals is especially sensitive to synchronization errors and relies on the capability of mobile devices to synchronize with base stations. This work proposes time and frequency offset estimation algorithms that minimize the residual synchronization errors in extreme coverage conditions. The proposed solutions were efficiently implemented with the resources of the developed modem SoC. Compared to other approaches commonly mentioned in the literature, the required acquisition time could be significantly reduced with improved estimation accuracy.

Cellular *OTDOA positioning* aims at expanding location-based services (LBS) to areas that are not covered by the GNSS, like indoor locations. The so-called time of arrival (TOA) measurements needed for OTDOA are the main challenges of its support in low-cost modems. In this thesis, approaches for TOA measurements in extended coverage were explored and implemented with the help of dedicated hardware acceleration. The impact of frequency offsets on measurement accuracy was also considered. The performance of the final solution was evaluated in a laboratory setup with a fabricated SoC.

The computational requirements of embedded modems are mainly dictated by the throughput and latency constraints imposed by cellular standards. Among the required signal processing, channel estimation and equalization are in the receive chain of most critical tasks. Thus, a *hardware accelerator for channel estimation and equalization* that meets the tight latency requirements of eMTC was designed. Time and frequency domain combination of channel estimations was included in the implementation to improve decoding performance in extended coverage. To minimize cost, hardware reuse within the block was extensively applied.

The implemented algorithms were evaluated with *measurements on a prototype modem* built around the developed SoC. The overall performance of the receive chain was assessed by considering the decoding performance of the main data channels. The impact of time-domain channel estimation averaging on eMTC sensitivity was analyzed and used to illustrate the importance of frequency synchronization. Furthermore, modem power consumption was analyzed, showing how sensitivity gains reduced energy usage when combined with early termination of data reception.

1.4 Outline

This thesis is organized as follows. Chap. 2 provides the background needed to understand the PHY of NB-IoT and eMTC. Chap. 3 presents algorithms to maintain fine time and frequency synchronization to the base station while operating in extended coverage. Chap. 4 introduces cellular OTDOA positioning and the signal processing solutions to support it in cIoT modems. Chap. 5 gives an overview of the architecture of the SoC jointly developed in this work and discusses the implementation of selected digital baseband processing tasks using a combination of embedded software and hardware acceleration. Chap. 6 presents measurements performed on the fabricated SoC, with considerations on synchronization and positioning performance, achievable sensitivity for communication, and power consumption. Finally, Chap. 7 concludes the thesis.

Chapter 2

Cellular IoT Background

Thanks to their global availability, cellular networks are commonly used for IoT applications that require wireless connectivity across large areas. For low-throughput use cases, 2G technology was often preferred to its 3G and 4G counterparts due to its lower complexity. However, to cope with the increasing number of connected devices and the rise of new low-power wide area networks (LPWAN), new alternatives were needed to improve the competitiveness of cellular networks for machine-type communications (MTC). Therefore, the 3GPP developed two standards based on LTE, namely NB-IoT and eMTC, to be part of its 5G portfolio for mMTC.

This chapter presents the physical layer (PHY) of NB-IoT and eMTC. The basic concepts needed to understand cellular communications are introduced first. Then, an overall picture of the two PHYs is given, before describing the specific features included to address the needs of MTC. Finally, the most challenging tasks that must be performed by the user terminals to fulfil the requirements of cIoT applications are identified.

2.1 Cellular Networks

Cellular networks are a kind of wireless access system where the covered region is divided into cells, each served by a base station (BS). The end devices, or user equipments (UEs), can communicate with each other or access the Internet through their serving BS. An ideal cellular grid with non-overlapping hexagonal cells is depicted in Fig. 2.1. In real deployments the covered areas can have different sizes and overlap. The UEs communicate with their serving BS with radio waves sent over the air, the information is then forwarded to a wired network that interconnects the BSs and allows the access to external networks. Signals from BSs to UEs build the downlink (DL) while those in the other direction the uplink (UL). Even though DL and UL transmissions are separated in time or frequency domain by time-division duplex (TDD) or frequency-division duplex (FDD), respectively, UEs can still suffer from interference from neighbour cells.

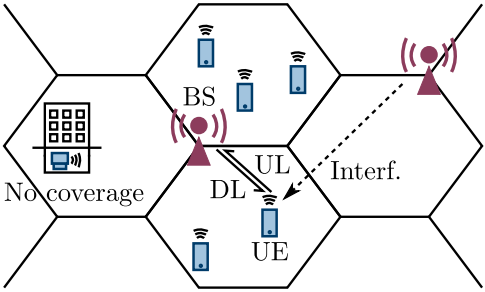

Figure 2.1: Example of a cellular grid.

The PHY of the radio access protocol defines how data packets are encoded and modulated to be transmitted through the wireless channel, setting, among others, the maximum throughput, the minimum latency, and the coverage that can be achieved. The real performance of a wireless network depends on how it is deployed and on the capabilities of the BSs and UEs that are used. The minimum requirements for the different components of the network are set by a

common standard. As cellular systems use licensed spectrum, efficient use of the expensive resources is of major interest for operators.

2.1.1 The Digital Communication Link

To achieve reliable communication and high efficiency, modern wireless communication standards use digital modulation techniques. In this case, the one-directional wireless link between two devices, e.g. the DL transmission from a BS to a UE, can be depicted as in Fig. 2.2.

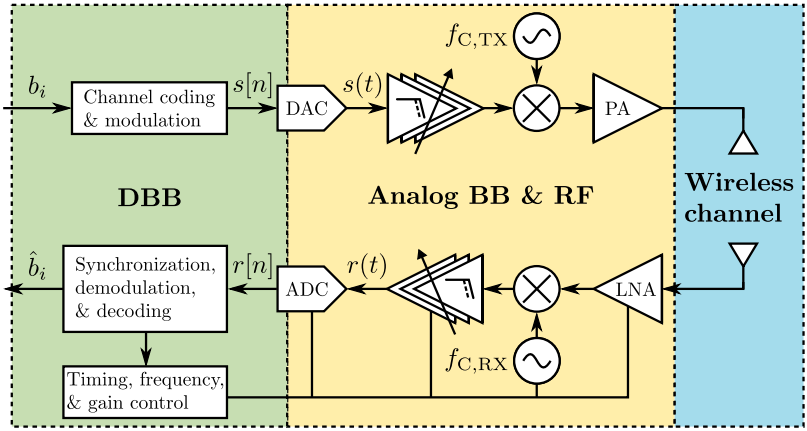

Figure 2.2: The wireless communication link.

On the transmitter side, the DBB signal generation includes the encoding of data bits b_i with forward error correction (FEC) codes, the modulation into data symbols, and the generation of the digital waveform $s[n]$ that is passed to the digital-to-analog converter (DAC). The analog signal $s(t)$ is then filtered and amplified before being up-converted to the RF f_C generated with a local oscillator as reference. Finally, the RF signal is amplified to the desired power level with a power amplifier (PA) and converted to a radio wave through an antenna.

	EPA	EVA	ETU
RMS delay spread	43 ns	357 ns	991 ns
Coherence bandwidth	4.5 MHz	460 kHz	100 kHz

Table 2.1: LTE channel models [16].

Wireless Channel Model

Besides suffering from interference, the radio wave is also subject to various undesired effects in its way from the transmitter to the receiver. One of the impacts of the radio channel is a power decay dependent on the distance travelled by the radio wave, its carrying medium, and the encountered obstacles. The larger the transmitted power and the smaller the needed receive power level for a successful transaction, the larger the covered area can be.

To model the effect of reflections and movement on the radio wave, the 3GPP used the LTE channel models for NB-IoT and eMTC. They are characterized by their delay profile, represented in the form of a tapped-delay-line, and by their Doppler spectrum, defined by the maximum Doppler frequency of the channel. Three multipath channel delay profiles are used with different average and maximum tap delays to cover various propagation models: extended pedestrian A (EPA), extended vehicular A (EVA), and extended typical urban (ETU) [16]. In the spectral domain, the multipath channels translate into a frequency-dependent scaling and rotation of the signal. The larger the delay spread of the channel, the faster its frequency response changes across the considered spectrum. The frequency range over which the effect of the channel can be considered to be almost constant is called the coherence bandwidth of the channel, it is inversely proportional to the time difference between the first and last tap with significant energy [17]. The root-mean-square (RMS) delay spread and the estimated coherence bandwidth for the different 3GPP channels are summarised in Tbl. 2.1. The coherence bandwidth was estimated by counting paths that are at most 10 dB below the strongest path.

Analog Receiver Model

The analog receiver also affects the received signal. The analog front-end and the low-noise amplifier (LNA) introduce thermal noise that can be modelled as additive white Gaussian noise (AWGN).

Furthermore, a mismatch between the carrier frequencies generated at the transmitter and the receiver causes a frequency offset $\nu = f_{\mathrm{C,TX}} - f_{\mathrm{C,RX}}$ in the down-converted baseband signal. The frequency offset can reach tens of kilo Hertz when the device is switched on for the first time and can vary strongly during operation due to the temperature-dependent behaviour of digitally-controlled crystal oscillators (DCXOs) commonly used for low-cost devices.

An inaccurate frequency reference also causes the time bases at the two ends to drift. Therefore, to maintain synchronization, the UE must periodically estimate its time offset (TO) to the BS and correct its time reference accordingly. The time offset can be modelled with a time delay τ in the received waveform.

Overall, the received baseband signal before sampling, including the mentioned effects of the analog receiver and the radio channel can be modelled as follows:

$$r(t) = Ae^{j2\pi\nu t}(s * h)(t - \tau) + n(t) \ . \tag{2.1}$$

The term $h(t)$ represents the channel impulse response and $*$ the convolution between two functions. For the 3GPP channel models that only have discrete taps h_i at specific time delays t_i, the impulse response is of the form:

$$h(t) = \sum_i h_i \delta(t - t_i) \ .$$

The amplitude scaling factor A includes the various amplification stages of the transmitter and the receiver, as well as the path loss due to radio transmission. To match the experienced noise power, the AWGN term $n(t)$ must be scaled to include the amplified noise introduced by the receiver.

Digital Signal Processing

The main goal of the digital signal processing in the receiver is to recover the transmitted bits from the noisy and distorted digital signal

$r[n] = r(t)|_{t=nT}$ delivered by the analog-to-digital converter (ADC). Besides handling the equalization, demodulation, and decoding of the received signal, the time and frequency synchronisation as well as receive (RX) gain control are also managed by the PHY in cellular communications.

While the higher protocol layers can be implemented in software, the tight throughput and latency constraints and the complexity of the PHY need dedicated hardware accelerators for DBB processing. On the other hand, minimizing silicon area by sharing computational resources across different tasks is highly desired in MTC devices to reduce cost. This work will focus on some of the digital signal processing tasks that are critical for coverage extension in cIoT communications and positioning, and on their implementation on low-cost SoCs.

2.2 The PHY of 5G Cellular LPWAN

Before discussing the challenges specific to cellular IoT systems, an overview of the PHY of the two standards considered in the scope of this work is given based on the more comprehensive description of [5]. Understanding the PHY helps specify the constraints that impose latency and computational requirements on the UE. Furthermore, it allows the analysis of the signals available to accomplish critical tasks like cell detection, synchronization, and equalization. The new challenges specific to MTC devices are identified highlighting the differences to legacy LTE.

2.2.1 Transmission Scheme and Numerology

The DL of NB-IoT and eMTC uses orthogonal frequency-division multiplexing (OFDM) with a subcarrier spacing of $f_\delta = 15\,\text{kHz}$. In the UL the two standards support single-carrier frequency-division multiple access (SC-FDMA) also with 15 kHz subcarrier spacing. In NB-IoT single-tone UL transmissions with either 15 kHz or 3.75 kHz tones are possible as well. The resources in the frequency domain are grouped into physical resource blocks (PRBs) containing 12 adjacent subcarriers. The system bandwidth for NB-IoT excluding guard bands is one PRB, i.e. 180 kHz. For eMTC, the UEs must support 1.08 MHz,

i.e. 6 PRBs or a so-called narrowband (NB), of RX and transmit (TX) bandwidth, but the cell can have up to 20 MHz bandwidth including guard bands. Instead of performing the OFDM modulation and demodulation at the full LTE sampling rate, eMTC UEs can tune their carrier frequency to the center of the NB assigned to their transmission and do the operations only on the targetted PRBs at 1.92 Msps. As for LTE, the time-domain structure of NB-IoT and eMTC is divided into frames of 10 ms, periodically numbered from 0 to 1023 and composed of 10 one-millisecond subframes (SFs) containing two slots. Each slot contains 7 OFDM symbols of approximately 71.4 µs for the normal cyclic prefix (CP), and 6 symbols of 83.3 µs each for the longer extended CP. NB-IoT supports only normal CP.

OFDM signals are typically depicted with a two-dimensional resource grid where time is represented on the horizontal axis and frequency on the vertical axis, as in Fig. 2.4. The smallest element of this grid corresponds to one subcarrier within an OFDM symbol and is called a resource element (RE) in 3GPP nomenclature. Their common numerology with LTE allows NB-IoT and eMTC to operate within LTE carriers in so-called in-band deployments without interfering with legacy transmissions. Only such in-band deployments are possible for eMTC while NB-IoT can also be deployed standalone or in the guard bands of LTE.

2.2.2 Physical Channels and Signals

All PHY transmissions can be categorized as physical channels or signals. Physical channels are used to convey user data as well as control information. The blocks of bits to be transmitted over physical channels are extended with cyclic redundancy checks (CRCs) for error detection and protected with FEC before being modulated on quadrature amplitude modulation (QAM) symbols and mapped to REs. Physical signals are pilots known at the transmitter and the receiver. They can be divided into synchronization signals and reference signals. The former are typically used for the initial cell search and the latter for channel estimation and equalization of data symbols.

DL physical channels and signals can be transmitted over multiple antennas using different antenna ports (APs). In the UL, eMTC

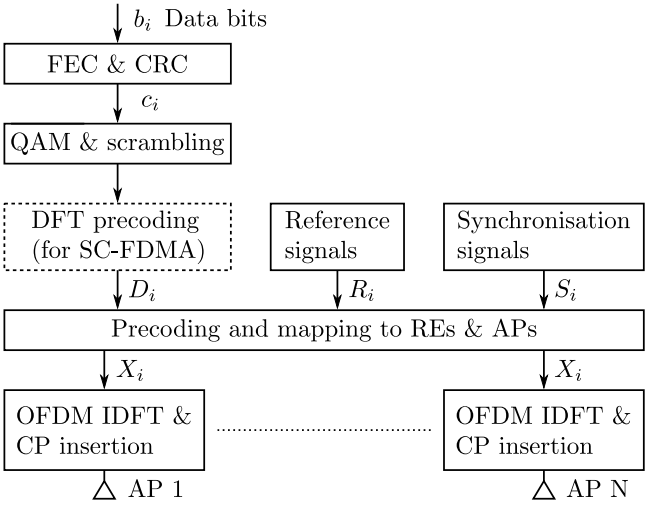

Figure 2.3: Baseband signal generation.

and NB-IoT UEs only support single-antenna transmissions. Physical transmissions can be assumed to experience the same channel only if they use the same AP [18]. On the one hand, this property dictates which reference signals must be used for the equalization of each data channel. On the other, it limits the signals that can be accumulated coherently to those belonging to the same AP. In some cases, this coherency condition is further limited in the time and frequency domains, allowing the BS to generate its signal with more flexibility. For example, AP hopping or PRB-specific beamforming could be used to better exploit the available degrees of freedom, but would at the same time restrict the number of subframes and subcarriers that can be combined coherently by the UE.

2.2.3 Overview of eMTC PHY

The DL OFDM resource grid of a 3 MHz FDD LTE cell supporting eMTC is shown in Fig. 2.4. The primary synchronization signal (PSS) and secondary synchronization signal (SSS) are the same as for LTE.

Both are periodically mapped every 5 subframes and occupy the 62 subcarriers around the center of the LTE cell's carrier frequency. The direct current (DC) subcarrier is left empty to prevent interference generated by leakage from the BS's carrier frequency. Always the same PSS out of three possible sequences, each corresponding to a different sector identity (ID), is mapped on the 7^{th} OFDM symbol of the SF. Two different SSS sequences are mapped on the 6^{th} OFDM symbol of SF 0 and SF 5. For each sector ID, 168 pairs of SSS sequences are possible, each corresponding to a group identity. The UE can determine the physical cell identity (PCID) of the LTE cell out of 504 possible values by combining the sector and the group identities after finding which sequences are mapped. Both PSS and SSS are sent on the same AP only if they are located in the same subframe, i.e. they cannot be assumed to be coherent across SFs. Besides PCID acquisition, PSS and SSS allow the detection of the CP type used in the cell, the acquisition of frame timing, and the initial frequency synchronization with the BS.

The relevant reference signals in the DL of eMTC are the cell-specific reference signals (CRS) and the demodulation reference signals (DMRS). CRS can be mapped on up to 4 APs. The number of BS antennas depends on the deployment and the UE must support all allowed configurations. In Fig. 2.4 they are depicted in blue for a cell with a single CRS AP. On a given AP, the CRS are transmitted coherently in each SF and across the full LTE cell bandwidth on every sixth subcarrier. For APs 0 and 1 they are mapped on the first and fifth symbol of each slot, and for APs 2 and 3 only on the second symbol. The CRS sequence and the subcarrier offset of the mapping in the frequency domain depend on the PCID, which is known only after PSS and SSS acquisition. CRS have a frequency reuse factor of six if a single AP is used and of three otherwise, as they are mapped every third subcarrier on symbols carrying APs 0 and 1. CRS can be used for time and frequency offset tracking because their presence does not depend on other transmissions. On the other hand, DMRS are always coupled to a physical channel and are only mapped if their associated channel is present. In the example of Fig. 2.4, they are only mapped in SFs 1 and 2. They are mainly used for the channel estimation needed to equalize the data symbols of physical channels.

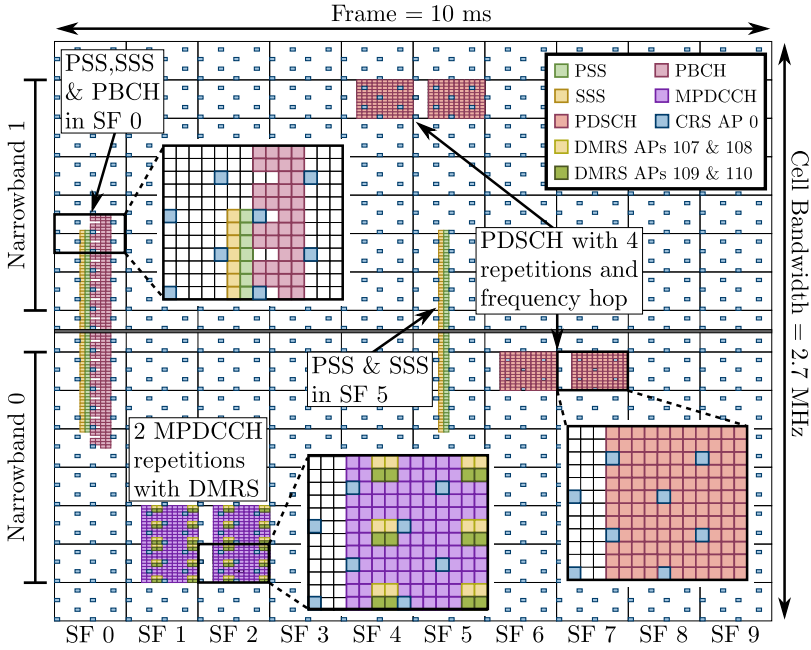

Figure 2.4: DL OFDM grid and resource allocation of eMTC in a 3 MHz LTE cell.

DMRS can only be assumed to be transmitted coherently within the same PRB and SF.

The first physical channel that has to be decoded by the UE is the physical broadcast channel (PBCH). It is mapped on the 6 central PRBs in the first SF of each frame and is transmitted on the same APs as the CRS. The PBCH is backwards compatible with legacy LTE and carries the so-called master information block (MIB) which contains basic information about the cell. After successfully decoding the PBCH, the UE can determine whether the cell supports eMTC and acquire essential knowledge on its configuration and status, like the DL system bandwidth, the up-to-date system frame number (SFN) and the number of CRS antenna ports. For cells supporting eMTC, the PBCH can be mapped with repetitions to extend the signal

coverage, one of the main goals of cellular LPWAN. However, the configuration of such repetitions is static and the decision to use them is up to the network operator. Thus, critical UE features should not rely only on PBCH repetitions.

As for legacy LTE, the default transmission duration of PHY channels, or transmission time interval (TTI), is 1 ms. The main PHY data channels are the physical downlink shared channel (PDSCH) and the physical uplink shared channel (PUSCH). They carry user data and higher-level control information in the DL and UL. Depending on the transmission mode, the PDSCH uses either CRS or DMRS as reference symbols. The empty REs in the first symbols of PDSCH SFs are reserved for the legacy control channel of LTE [19]. For eMTC, PDSCH and PUSCH had to be limited to occupy between one and six adjacent PRBs to enable UEs with reduced RX and TX bandwidths. Both use turbo coding for FEC and can be modulated with quadrature phase-shift keying (QPSK) or 16-QAM symbols. The rate used for channel coding can be adjusted by puncturing part of the bits generated by the FEC encoder. The selected combination of coding rate and mapped QAM symbols, called the modulation and coding scheme (MCS), is selected by the BS. For transmissions dedicated to a specific UE, the MCS can be adapted to fit the conditions of the radio link. For broadcast data, the configuration is static and should be set by the operator depending on the area that must be covered by a cell. To deal with the reduced sensitivity of low-cost single-antenna UEs and to achieve the coverage extension goals that will be discussed later in this chapter, eMTC allows all channels to be mapped with repetitions, expanding the transmission across multiple TTIs. The maximum number of repetitions that can be used for PDSCH and PUSCH depends on the coverage enhancement (CE) mode in which a device operates. CE mode A is mandatory and allows up to 32 repetitions for both channels, while CE mode B allows up to 2048 repetitions but its support is optional for the UE and the BS. Like the MCS, the number of repetitions of data channels can be dynamically adjusted by the BS.

Control information for the PHY is exchanged through dedicated channels. The MTC physical downlink control channel (MPDCCH) carries downlink control information (DCI) from the BS to the UE. It includes scheduling and coding information needed by the end device

to receive PDSCH and send PUSCH transmissions. In the UL, PHY control information like the estimated strength and quality of the signal received by the UE is conveyed through the physical uplink control channel (PUCCH). These two control channels are also used to report hybrid automatic repeat request (HARQ) feedback for data channels in both directions, DL and UL.

The time relationship between control channels and their associated data channels is imposed by the standard and they set hard constraints on the latency of DL signal processing and UL signal generation. As depicted in Fig. 2.4, the UE has 1 ms between the end of the MPDCCH and the beginning of the PDSCH that it schedules. Since the narrowband used for the transmission of the PDSCH is part of the DCI carried by the MPDCCH, the UE has to successfully decode the MPDCCH and extract the scheduling information early enough to retune its frequency to the desired narrowband for the reception of the PDSCH. A similar constraint arises from the timing of UL scheduling. In this case, with the maximum timing advance (TA) applied to UL channels to compensate for the round-trip delay of the radio waves, the end of the MPDCCH and the beginning of the scheduled PUSCH are separated by 2.3 ms at least. In this time interval, the PHY of the UE must not only decode the DCI but also fetch the UL data from higher protocol layers, generate the PUSCH, and start the RF transmission. The latency requirement for PDSCH decoding is set by the timing of the associated PUCCH for HARQ feedback. In this case, the minimum delay to decode the PDSCH and start the transmission of the PUCCH is again 2.3 ms.

2.2.4 Overview of NB-IoT PHY

NB-IoT has a set of physical signals and channels similar to eMTC, but only supports up to two APs in the DL. An overview of possible DL transmissions is shown in Fig. 2.5. The signals available for initial cell search and synchronization are the narrowband PSS (NPSS) and the narrowband SSS (NSSS). To compensate for the reduced signal bandwidth of NB-IoT, the synchronization signals have a longer duration than their LTE and eMTC counterparts. NPSS and NSSS both occupy 11 OFDM symbols, leaving the three first symbols of the SF unused to avoid collisions with legacy LTE in in-band deployments.

The former is mapped on 11 out of 12 subcarriers in subframe 5 of each frame, while the latter occupies all 12 subcarriers of the last SF in all even frames, i.e. with a periodicity of 20 ms. As opposed to the three possible PSS sequences in LTE, only one NPSS is available in NB-IoT. A single NPSS sequence has the advantage of reducing the complexity of initial cell search and synchronization by limiting the search space that has to be covered. However, it comes at the cost of less flexibility for NPSS interference mitigation in network planning, because all cells have to use the same signal. The PCID is fully determined by the NSSS sequence. Like in LTE, NPSS and NSSS can be assumed to be sent on the same antenna port only within a subframe.

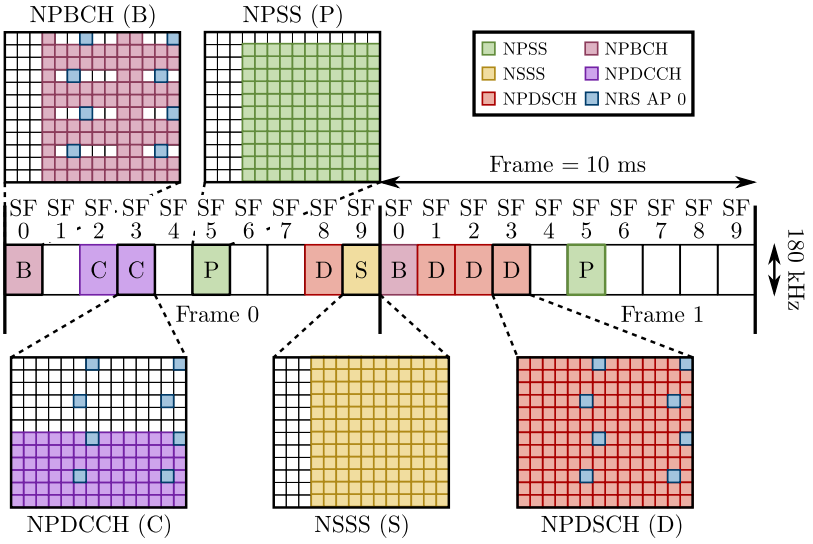

Figure 2.5: DL OFDM grid and resource allocation of an NB-IoT cell.

All NB-IoT downlink channels are associated with the same set of pilots, the narrowband reference signals (NRS). These reference signals can be sent on one or two antenna ports and are mapped on two subcarriers of the last and second-last symbols of each slot, as shown by the blue squares in the detailed resource grids of Fig. 2.5.

The occupied subcarriers and the used sequence of QPSK symbols depend on the PCID. NRS have a frequency reuse factor of six for one AP, and three for two APs. Overall, they occupy 8 REs per antenna port on the SFs where they are present. Unlike CRS, NRS are only sent on a subset of subframes defined by the cell's configuration. The small number of SFs where NRS can be assumed to be always present complicates the use of these pilot signals for radio link monitoring and synchronization at a low signal-to-noise ratio (SNR).

The main broadcast channel, narrowband PBCH (NPBCH), as well as the downlink control and data channels, narrowband PDCCH (NPDCCH) and narrowband PDSCH (NPDSCH), are also depicted in Fig. 2.5. They fulfil similar functions as their eMTC counterparts. All DL physical channels rely on NRS for channel equalization and a convolutional code for FEC. In the UL, narrowband PUSCH (NPUSCH) is used to carry user and higher layer control data as well as HARQ feedback. For data transmission, NPUSCH is turbo encoded, while for HARQ information repetition coding is used. To support large payload block sizes despite a reduced signal bandwidth, data channels can span across multiple SFs even without repetitions. Thus, NB-IoT transmissions tend to have a longer duration and lower throughput compared to eMTC. All DL and UL channels also support repetitions with up to 2048 NPDSCH and NPUSCH repetitions mandatorily supported by all UEs.

Besides lower peak data rates, NB-IoT also has higher delays between control and data channels, relaxing the latency requirements of PHY processing. In the UL, the minimum delay between the end of the NPDCCH and its scheduled NPUSCH is 7.3 ms. As for the DL, the delay between control and data channels is 4 ms, and the minimum gap between the end of an NPDSCH and the start of its associated HARQ NPUSCH is 11.3 ms.

2.3 Requirements of Cellular IoT

NB-IoT and eMTC were introduced to tackle the needs of cIoT applications. The main requirements that drove the development of 3GPP's standards for MTC highlight the challenges faced in the design of competitive UE modems that must deal with contradicting goals.

2.3.1 Cost Reduction

One of the obstacles to using LTE to connect a massive number of devices is the price of state-of-the-art modules common in high-end mobile broadband applications. The minimum set of features that must be supported by legacy LTE UEs already requires costly hardware components that are prohibitive for applications with high cost-pressure that only need small data rates. Cellular IoT standards allow a considerable reduction in device price by enabling low-complexity terminals [20, 21].

Limiting the maximum RX and TX bandwidths reduces the complexity of analog RF and baseband processing blocks at the cost of smaller achievable peak data rates and reduced frequency diversity. It also enables lower sampling rates and cheaper ADC and digital front-end (DFE) designs. It can also simplify many DBB processing tasks such as the fast Fourier transform (FFT) used for OFDM modulation and demodulation. In Tbl. 2.2 the minimum RF bandwidth, sampling rate, and OFDM FFT size of NB-IoT and eMTC are compared with legacy LTE devices.

Furthermore, the smaller size of data packets in the targetted mMTC use cases allows the reduction of the maximum block lengths supported by the PHY. The shorter blocks combined with the lower data rates result in relaxed requirements for the channel decoders, allowing smaller implementations. Increasing the delay between control and data channels for scheduling as well as for HARQ feedback also reduces the latency requirement of the RX and TX chains, further facilitating the complexity reduction of channel decoders. This is

	LTE	eMTC	NB-IoT
RF bandwidth	20 MHz	1.4 MHz	200 kHz
Minimum sampling rate	30.72 Msps	1.92 Msps	240 ksps
OFDM FFT size	2048	128	16
Peak data rate DL/UL	10/5 Mbps[1]	300/375 kbps	25/62 kbps

Table 2.2: Comparison of bandwidth parameters.

[1]Maximum PHY throughput for LTE Category 1 UEs.

particularly valid for NB-IoT, where at least 7.3 ms are available for UL scheduling, and 11.3 ms for DL HARQ feedback. In the DL of NB-IoT, data decoding was simplified further by using a convolutional code instead of a turbo code for NPDSCH forward error correction (FEC). For eMTC, the real-time requirement is more challenging, especially for UL scheduling, because the minimum time between an MPDCCH carrying an uplink schedule and the corresponding PUSCH is approximately 2.3 ms. In that period, the PHY must decode all MPDCCH candidates, fetch the data to be transmitted from higher layers, and generate the PUSCH.

Enabling half-duplex FDD (HD-FDD) UEs with at least 1 ms to switch from DL to UL and vice versa eliminates the need for a duplexer and for separate local oscillators to generate RX and TX frequencies. Another common choice made by module vendors to reduce device cost is to use DCXOs instead of more expensive temperature compensated crystal oscillators (TCXOs) as frequency references. However, using less accurate DCXOs doesn't only require robust tracking algorithms for periodic frequency correction but also drastically increases the frequency range that must be covered during initial synchronization.

Both NB-IoT and eMTC UEs can operate with a single antenna and save costs and space by avoiding multiple RX chains. The resulting losses of array gain and spatial diversity in the receiver translate into a reduced DL coverage that must be compensated.

Overall a cost reduction of approximately $60 - 70\,\%$ for eMTC HD-FDD modems [5] compared to legacy LTE UEs should be achieved by exploiting the relaxed requirements and the complexity reduction techniques. For NB-IoT even lower complexity UEs are possible due to its smaller bandwidth, lower throughput, simpler decoding, and higher latency.

2.3.2 Low Power Consumption

In many applications the use of battery-powered systems is driven by mobility requirements, but also for static devices, such as utility meters, it may be impossible or too expensive to use the main grid as a power source. The two use cases benefit from low-power operation either allowing smaller batteries or increasing the lifetime of a single

charge. To minimize maintenance costs and make install-and-forget deployment models possible for products in remote areas, at least 10 years of battery life should be possible with a standard 5 W h battery assuming infrequent transmission of small data packets [5].

The solutions introduced to optimize power consumption are extended discontinuous reception (eDRX) and power saving mode (PSM). The former builds on top of legacy discontinuous reception (DRX) by allowing longer inactivity periods. Once per eDRX cycle, the UE should perform measurements and monitor control channels during paging transmission windows containing at least one DRX cycle. The latter allows the UE to completely skip measurements and to monitor paging only during short periods after mobile-originated data transfers or tracking area updates (TAUs), both involving uplink transmissions. For PSM the TAU periodicity defines how often the UE is guaranteed to be reachable while for eDRX it is the eDRX cycle length. If the inactivity periods are long the UE has to synchronize again to the BS when it wakes up.

As reachability comes at the cost of power-hungry TAUs in PSM, this mode is intended for scenarios where devices can be unattainable for long periods and data transfers are mainly initiated by the UE, as for utility meters. In PSM, the main contributors to energy consumption are the deep sleep current and the TX power of UL transmissions. In eDRX, the power for paging monitoring and neighbour cell measurements can become dominant for short eDRX cycles. For long inactivity periods combined with extended coverage operation, cell resynchronization after wake-up can also significantly impact battery life because the UE will have to receive multiple SFs to accumulate enough power to accurately estimate the time and frequency offsets.

Overall, when the device operates at low SNR, good receiver performance can benefit energy efficiency by reducing the number of RX SFs needed to decode DL control or data channels. The same applies to cell search and time and frequency synchronization, if the UE can perform them with a shorter received signal using efficient algorithms, it will save power by switching off its receiver earlier than with less performant algorithms. Tracking and measurements are important because they have to be periodically done by the UE not only to set up and maintain a connection but also to monitor paging after a long idle time and measure neighbouring cells for reselection.

2.3.3 Coverage Extension

Coverage extension is another important feature of new cellular IoT standards. Not only do the losses introduced by some of the complexity reduction techniques have to be compensated for, but the coverage should also be significantly extended compared to legacy LTE and Global System for Mobile Communications (GSM) devices to enable deep indoor applications such as smart metering and building monitoring where the devices may have to operate underground.

The coverage is measured as the maximum coupling loss (MCL) at which a connection can be maintained with sufficient performance, i.e. the UE should stay synchronized to the cell, and DL and UL physical channels should operate with a reasonable block error rate (BLER), e.g. $5 - 10\,\%$. The MCL of a physical channel is the difference between the TX power across the spectrum occupied by the channel, and the sensitivity of the receiver, which corresponds to the minimum RX power level at which the targeted BLER can be achieved :

$$\text{MCL [dB]} = \text{TX Power [dBm]} - \text{Sensitivity [dBm]} \; . \tag{2.2}$$

The sesitivity itself depends on the received signal bandwidth, the noise figure (NF) of the analog receiver and the SNR required to decode the channel with that BLER [5]:

$$\begin{aligned} \text{Sensitivity [dBm]} = {} & \text{Thermal Noise [dBm/Hz]} \\ & + 10 \log_{10} (\text{Channel Bandwidth [Hz]}) \\ & + \text{Receiver NF [dB]} \\ & + \text{Required SNR [dB]} \; . \end{aligned} \tag{2.3}$$

The 3GPP agreed on a target MCL of 164 dB to achieve a 20 dB coverage enhancement over legacy cellular standards. The improvement can come from the different terms that make up the MCL in (2.2) and (2.3). One option would be to rely on reduced NFs in analog receivers. However, modern devices only have a small margin for improvement as they already have noise figures in the order of 3 dB for base stations and 5 dB for UEs.

Another possibility to improve the MCL would be to increase the overall TX power. But the maximum transmitted power is limited to fulfil radiation regulations and to reduce PA complexity and peak

power consumption. Instead, power spectral density (PSD) boosting can be used. In the DL, this can be done by allocating more power to the spectrum used by devices in poor coverage while keeping the total power constant. For eMTC, a PSD boosting of up to 4 dB should be possible in general [22] while for NB-IoT the BSs should support at least 6 dB boosting [23] for in-band and guard-band deployments. For standalone NB-IoT cells, no boosting is needed as the TX power of the BS is already concentrated within 200 kHz. In the UL, PSD boosting can be done by reducing the signal bandwidth for the same TX power. The minimum transmitter bandwidth of eMTC UEs corresponds to one PRB and does not provide any gain compared to legacy LTE. On the other hand, NB-IoT supports UL transmission bandwidths as low as 3.75 kHz with a better trade-off between energy consumption and system capacity for extended coverage scenarios and small data rates [24].

Two mechanisms that are available in legacy cellular standards to reduce the SNR at which data can be decoded: flexible MCS and HARQ repetitions. The former is already well exploited in LTE and only has limited potential for improvements, while the latter is only suited for a small number of repetitions due to the overhead of HARQ signalling. To significantly increase coverage compared to LTE and GSM, automatic repetitions were introduced for all PHY channels. The number of repetitions is controlled by the BS. For data channels, it is dynamically set in the DCI, and for control channels, it is part of common or UE-specific configurations. At most 2048 repetitions are supported by both NB-IoT and eMTC.

Ideally, the coverage should improve by 3 dB per doubling of the number of repetitions. With that assumption, up to 33 dB could be gained compared to the single-repetition case. However, repetitions yield diminishing returns with decreasing SNR because the performance is then limited by the accuracy of the channel estimation and the residual time and frequency offsets. To achieve the targeted MCL without DL PSD boosting, the UE must operate down to -18.5 dB SNR [5]. Under these conditions, the coverage of a UE is limited by the challenging tasks implicitly needed to set up and maintain a connection, like cell search and time and frequency synchronization.

2.3.4 Positioning

Positioning capabilities are the basis for a wide range of applications covered by so-called location-based services (LBS). Typical massive IoT (mIoT) use cases that require an estimation of the terminal's position span from smart wearables to emergency call systems and asset trackers for logistics or delivery services. Depending on the scenario, the required location accuracy can range from a few meters to hundreds of meters. Thanks to its global availability and relatively high precision, GNSS has established itself as the preferred positioning technology for mainstream applications like automotive navigation and smartphones. Thus, GNSS is also a natural choice for many MTC applications.

However, applications that require indoor coverage cannot solely rely on GNSS. To cover areas where satellite signals cannot penetrate, Bluetooth and wireless local area network (WLAN) deployments can be used as alternatives. Especially, Bluetooth Low Energy (BLE), with its low device power consumption and cost, has attracted increasing interest as an option for indoor positioning solutions based on constellations of low complexity beacons [25]. However, the infrastructure needed for LBS with BLE is not globally accessible and would have to be deployed locally where needed, inducing higher costs than approaches based on already available networks.

Another option is to take advantage of the existing cellular framework and employ it for positioning. For instance, assisted GNSS (A-GNSS), where the cellular network is used to deliver assistance data to the UE, helps reduce the acquisition time and power consumption of GNSS positioning [6]. Furthermore, A-GNSS can be integrated into cIoT modems at negligible cost, if the cellular receiver and other available resources are reused efficiently [26]. Additionally, standalone cellular positioning is also possible without the need for a separate technology, thus avoiding hardware (HW) overhead like dedicated antennas or additional modules. For example, enhanced cell ID (E-CID) is a technique used in LTE that can be easily transposed to eMTC and NB-IoT. It relies on measurements performed by the UE and the BS to refine an estimation of the UE's location based on the position of its serving BS [27, Chap. 32]. However, the location accuracy of E-CID is not always sufficient. Therefore, the 3GPP

introduced observed time difference of arrival (OTDOA) positioning in LTE Release 9, with the original goal of complying with the requirements of modern emergency services [28]. Even though it is not as precise as GNSS, OTDOA positioning can also cover the needs of a wide range of MTC applications, hence the effort made by the 3GPP to include OTDOA support in the second release of eMTC and NB-IoT. OTDOA relies on so-called reference signal time difference (RSTD) measurements performed by the UE on signals from different base stations. The implementation of precise time delay measurements with the processing resources available on a low-cost system is the main challenge faced to support OTDOA in cIoT modems.

2.4 User Equipment Challenges

The mentioned requirements of mMTC applications modify the challenges of the development of cIoT terminals compared to typical mobile broadband devices. While the latter require high throughput decoders and complex multiple-input multiple-output (MIMO) processing to achieve the highest data rates supported by the standard, the main challenges of mMTC devices are to operate at extremely low SNRs and to offer sufficient communication and positioning capabilities while minimizing modem complexity and power consumption.

Indeed, specially tailored algorithms are needed to fully exploit repetitions when the wanted signal is well below the noise floor, making coverage extension the most challenging MTC-specific goal for the UE. Among the different signal processing tasks that need to be adapted to operate far below the noise floor, techniques for fine time and frequency offset tracking, time delay measurement for OTDOA positioning, and channel estimation for the equalization of received signals will be discussed in the scope of this work. The preferred methods will be analyzed and selected according to their performance at low SNR and based on their suitability for implementations on low-cost and highly integrated modem SoCs. Thus, special attention will be focused on the system architecture and the execution of the discussed algorithms with the available resources.

2.4.1 Fine Synchronization

Accurate synchronization is of utmost importance for both positioning and wireless communications. Like GNSS, the accuracy of cellular OTDOA positioning depends on the precision of the time delay measurements performed by the end device, which corresponds to its ability to synchronize in time with the received signals. In this case, residual frequency offsets degrade the performance of the measurement, also calling for accurately matched frequency. As for cellular communications, even though synchronization is implicitly needed to decode DL data channels, it is also essential to limit the interference between the UL transmissions of different UEs. In 3GPP's cellular standards, the minimum synchronization requirements for UEs are defined and verified in the UL. For eMTC and NB-IoT, the frequency and timing of the UE must track the base station's reference within approximately 100 Hz and less than a few microseconds [16, 29].

Although the synchronization requirements of eMTC and NB-IoT are close to the ones of legacy LTE, or moderately less stringent in some cases, the characteristics of cellular IoT communications impose other challenges. To minimize cost, DCXOs are typically preferred to more expensive and accurate TCXOs for cIoT modems. The higher vulnerability to temperature variations of such references calls for more frequent and larger timing and frequency corrections. Furthermore, to perform these adjustments in extended coverage scenarios, the UE must be able to estimate the residual frequency offset of its local oscillator at extremely low SNRs. Thus, performant algorithms that fully exploit the available pilot signals are needed to allow tracking in extended coverage and the minimization of the required measurement accumulation time to limit power consumption. Furthermore, residual frequency offsets affect the coherency of the received signal over time, limiting the SNR gain that can be achieved with repetitions of physical channels. Therefore, it is advantageous to keep the residual frequency offset as low as possible, well below the standard requirements, to optimize data transmission in extended coverage to allow even lower SNRs or power savings by reducing the required accumulation time. These considerations highlight the need for precise time and frequency offset estimation algorithms with high

accumulation gain and sufficient estimation range, a major topic of this work.

2.4.2 Channel Estimation and Equalization

Coverage extension is also one of the main challenges of data transmission in cellular IoT communications. For the UE, it implies the need to decode DL channels at least down to $-18.5\,\mathrm{dB}$ SNR. As already mentioned, the mechanism used to counter highly negative SNR is to automatically repeat data channels. The repetitions are commonly combined by adding the soft log-likelihood ratios (LLRs) obtained from the equalized symbols before channel decoding [30]. Even though the combination can also take place before the equalizer for a limited amount of repetitions [31], the coherent modulation used for NB-IoT and eMTC requires the equalization of received symbols at some point to restore the phase information. The quality of the equalized symbols is highly dependent on the channel estimation which becomes one of the limiting factors of the decoding performance. Furthermore, when the repetitions are combined in the LLR buffer, the accumulation gain is highly dependent on the coherency of the equalized symbols which, again, is limited by the channel estimation.

Besides the need for accurate channel estimations for coverage extension, the limited memory space and processing capabilities of cIoT UEs also need to be considered in the choice of the used algorithms. Moreover, channel estimation and equalization have stringent latency requirements because they are part of the signal processing path for the scheduling of DL and UL data channels and the transmission of DL HARQ feedback. A solution that seeks the right compromise between complexity and channel estimation quality will be discussed in this work.

2.4.3 OTDOA Positioning

Either as a low-cost standalone solution or as a fallback of GNSS for indoor scenarios, OTDOA positioning is a valuable feature for cIoT devices. Its precision relies mainly on the accuracy of the RSTD measurements used to estimate the distance difference between the UE and different BSs. To allow indoor applications, the measurements should

work well below the noise floor, with standard requirements specified down to −15 dB SNR. Furthermore, the reduced signal bandwidth of eMTC and, especially, NB-IoT deteriorate the precision of timing measurements. Thus long signal accumulation periods are needed to cope with the small bandwidth and the targeted coverage regime.

Thus, performant implementation of the RSTD measurements with the available pilot signals is crucial to support OTDOA positioning on low-cost UEs. Furthermore, to limit the memory required to combine many signal periods, the accumulation needs to be computed on-the-fly. A measurement strategy, as well as its implementation with hardware acceleration, will be presented in the scope of this work.

2.4.4 System Design and Hardware Acceleration

The variety of applications and the heterogeneous adoption of emerging standards across the world, combined with the high pressure on cost and form factor for mIoT devices, make the design of one-fits-all solutions almost impossible. In this context, integrated multimode UEs are most likely to provide truely global coverage at minimum cost and size. Considering mIoT scenarios with relaxed latency requirements and long idle periods, it is possible to support multiple standards on the same system with minimum overhead by sharing HW components. For instance, a cellular modem and a GNSS receiver could share the same processing system and part of the analog receive chain [26].

The same applies to the PHY processing for cellular communications, where computing and memory resources can be shared for different standards. However, dedicated hardware accelerators are needed to compensate for the limited capabilities of low-power processing systems and meet the tight latency constraints of cellular standards. Furthermore, performing computational intensive tasks on dedicated HW also provides an advantage in terms of power efficiency compared to a software (SW) implementation. Special attention will be devoted to the architecture that was developed to integrate various communication and positioning technologies into a single SoC, highlighting the chosen partitioning between HW and SW for the implementation of time and frequency offset tracking, channel estimation and equalization, and OTDOA positioning.

Chapter 3

Fine Synchronization in Extended Coverage

Accurate synchronization is essential to wireless communication and positioning systems. Widely used OFDM signals suffer particularly from time and frequency mismatches between transmitter and receiver because synchronization errors introduce inter-symbol interference (ISI) and inter-carrier interference (ICI) [32]. Furthermore, time and frequency offsets tend to be worse in low-cost IoT modems, compared to high-end terminals, due to the highly temperature-dependent frequency of DCXOs commonly used for such devices. Additionally, any residual frequency offset (FO) limits the gain of coherent accumulation over long periods, needed for coverage extension. Hence, to achieve coverage extension in cIoT networks, end devices must minimize their residual FO to optimize accumulation gains. When operating below the noise floor, the coherency loss has typically a higher impact on the link performance than the aforementioned ISI and ICI terms. Maintaining accurate synchronization in coverage extension is itself a challenging task that has only recently gained attention in the literature [15,33] with the emergence of cIoT standards and needs specially tailored algorithms to fully exploit the available resources for time and frequency offset estimation.

This chapter describes the general problem of fine synchronization and presents practical solutions that can be applied in cIoT systems. First, the operation of the tracking loop in the considered system is explained and the estimation of time and frequency offsets is highlighted as the limiting factor of synchronization performance. Then, an exact model of the effect of time and frequency offsets on the demodulated OFDM signal is derived. Following, the Cramer-Rao lower bound (CRLB) of the achievable estimation accuracy in AWGN channels is investigated. The rest of the chapter is devoted to the investigation of different algorithms for fine synchronization. The proposed solutions are discussed in the context of eMTC, but they are also applicable for coverage extension in other OFDM systems like NB-IoT and 5G NR.

3.1 Synchronization in Cellular Systems

Wireless data transmission typically relies on some kind of synchronization between the two ends of a communication link. In cellular networks, mobile terminals must synchronize their local time and frequency references to BSs to be able to decode DL data and transmit UL signals without interfering with other devices. This is commonly done in different steps, starting with a coarse synchronization followed by one or more fine correction stages. Not only do the residual offsets after initial coarse acquisition have to be corrected, but synchronization must be maintained as long as the UE is connected or is actively monitoring DL channels. For cIoT devices in low to medium mobility applications, frequency offsets are mainly caused by inaccurate references and can be directly corrected by recalibrating their local oscillators. Time offsets can be caused by changes in the path delay if the terminal is moving, but they can also gather for static scenarios due to residual sampling frequency offsets. Especially, low-power UEs typically operate with low duty cycles, i.e. with long inactivity periods between short activity bursts. Examples of such operating modes are eDRX and PSM, already described in Chap. 2. In these configurations, the UE must resynchronize after waking up from idle cycles to be ready for active phases. Long sleep times allow time offsets to build up even with small residual sampling offsets.

A UE structure as depicted in Fig. 3.1 is considered, where the carrier frequency used to down-convert the radio signal, the sampling clock of the ADC, and the time-keeping clock are derived from the same frequency reference. A tracking loop is built by estimating and correcting the synchronization mismatches of the received signal.

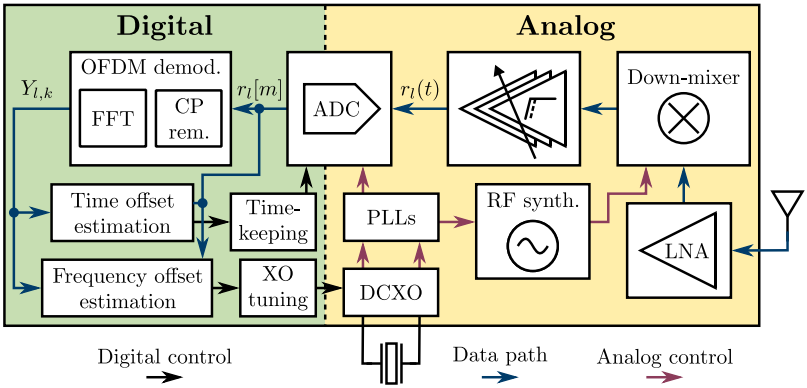

Figure 3.1: Illustration of considered receiver chain with time and frequency tracking loop.

The residual FO is corrected directly by tuning the DCXO. Thus, carrier, sampling, and time-keeping frequencies are adjusted together. As long as the range and the granularity of the DCXO tuning are sufficient, the residual frequency offset is not impacted by the correction step. Time offsets are compensated by adapting the starting time of the ADC, or by dropping or padding samples.

The most important step to ensure small residual impairments is the estimation of the TO and FO that have to be corrected. To estimate these offsets, the UE relies on the signals periodically transmitted by the BS. The choice of the estimation method and how the UE exploits the available pilots impact the performance of the system. The computational requirements of certain algorithms can be prohibitive for low-cost devices. On the other hand, complex algorithms may need less accumulation time to reach the target accuracy, limiting the required RX time of the transceiver and therefore reducing the power consumption of the terminal. Finally, minimizing the residual

synchronization errors improves the coherency of the received signals and allows a better usage of the repetitions of data transmissions. Thus, accurate estimation of time and frequency offsets at low SNR is crucial for coverage extension.

3.2 Effects of Time and Frequency Offsets

Understanding the effects of synchronization mismatches on the received digital baseband signal and on the elements of the OFDM resource grid after demodulation is crucial to developing time and frequency offset estimation algorithms. A TO in the local time reference of the UE results in a misalignment between the receive window and the actual time window of the OFDM symbol. In [34] the TO is shown to introduce ISI, ICI, and a rotation of the demodulated OFDM symbols with a phase proportional to the subcarrier index. The ISI caused by the TO results from the overlap of the shifted receive window of the l^{th} OFDM symbol with the next symbol, as depicted in Fig. 3.2. The subcarrier-dependent phase shift introduced by the TO is also highlighted in the figure. It can be visualized as the changing phases at the beginning of the receive window on the sinusoid of each subcarrier, marked with coloured crosses.

A residual offset in the local oscillator's frequency causes both a carrier frequency offset (CFO) and a sampling frequency offset (SFO). As the two frequencies are derived from the same reference, the offsets are linearly dependent in the considered system. The SFO can be modelled as an offset in the sampling period, while the CFO results in an FO in the down-mixed baseband signal. A detailed derivation of the joint effects of small SFO and CFO on OFDM signals can be found in [35], for the general case where they are independent. The introduced ICI and phase rotation across OFDM symbols are shown to depend on both, the CFO and the SFO, while the rotation across subcarriers depends only on the SFO. The ICI due to a CFO is illustrated on the left-hand side of Fig. 3.2, where the received subcarriers are plotted in the frequency domain. When subcarrier k is demodulated with the correct frequency, it is sampled where the contribution of all other subcarriers is equal to zero. On the other hand, when the k^{th} subcarrier is sampled with a CFO, as shown by

the dotted blue arrow, ICI from other subcarriers is introduced. The contribution of the neighbouring subcarriers to the ICI is highlighted on the left plot of the figure with coloured dots.

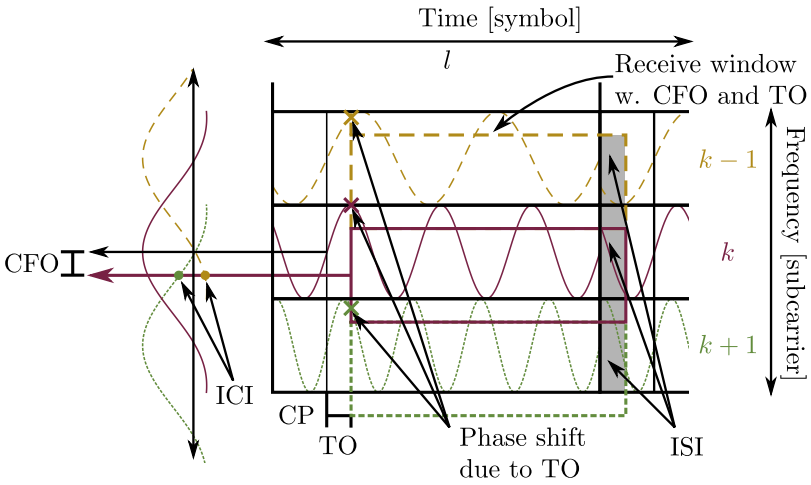

Figure 3.2: Visualization of the effects of TO and CFO on OFDM signals.

ICI and ISI only have a limited impact on the system when the UE is tracking and the residual offsets are kept small. In most scenarios, the performance is mainly limited by noise or interference. Concerning the effects of synchronization mismatches, the focus will be set on the phase rotations across subcarriers and OFDM symbols. These are relevant because they can be used to estimate time and frequency offsets and because they limit the coherency of the signal. The exact influence of the CFO and TO on the demodulated REs will be investigated to derive a quantitative formula for the phase rotations. The results will then be compared with the effect of the SFO investigated in previous works to show that the SFO can be neglected.

3.2.1 Mathematical Model

A mathematical model for the effects of the CFO and TO on the OFDM demodulated symbols can be derived starting with the formula of the transmitted waveform and then following the same steps as in [35]. The continuous-time signal for the l^{th} OFDM symbol is generated as follows:

$$s_l(t) = \frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{k=-\frac{N_{\text{FFT}}}{2}}^{\frac{N_{\text{FFT}}}{2}-1} X_{k,l} e^{j2\pi f_\delta k(t-T_{\text{CP}}-lT_{\text{SB}})} \ . \tag{3.1}$$

$X_{k,l}$ is the QAM symbol that modulates the k^{th} subcarrier. The parameter T_{CP} is the duration of the cyclic prefix, T_{SB} is the duration of one OFDM symbol including the CP, and f_δ is the subcarrier spacing. (3.1) holds for $lT_{\text{SB}} \leq t < (l+1)T_{\text{SB}}$, while $s_l(t) = 0$ otherwise.

The digital OFDM signal before analog conversion is usually generated with an inverse fast Fourier transform (IFFT) of size N_{FFT} at a sampling frequency $f_s = f_\delta N_{\text{FFT}}$. Guard bands at the edge of the signal's spectrum are normally included by mapping zeros to the corresponding subcarriers.

After upconversion to the carrier frequency, transmission through the radio channel with impulse response $h_l(t')$, downconversion back to baseband with a CFO of ν Hertz, and delay by a TO of τ seconds, the received baseband signal is:

$$r_l(t) = e^{j2\pi\nu t}(s_l * h_l)(t-\tau) + n(t) \ . \tag{3.2}$$

The symbol received at the k^{th} subcarrier after sampling at frequency f_s, i.e. $r_l[m] = r_l(t)|_{t=\frac{m}{f_s}}$, and OFDM demodulation with an FFT of size N_{FFT} can be approximated as in [35] for small offsets[1]:

$$Y_{k,l} = X_{k,l} H_{k,l} e^{j\phi(k,l,\nu,\tau)} + W(k,l,\nu,\tau) + N_{k,l} \ . \tag{3.3}$$

$N_{k,l}$ is a noise term resulting from $n(t)$ in (3.2) after sampling and OFDM demodulation. $W(k,l,\nu,\tau)$ is the ICI introduced by the CFO and can be neglected as long as the FO is small compared to the

[1] Derivations can be found in App. A.1.

subcarrier spacing f_δ. The exact formula of the ICI term can be found in App. A. The term $H_{k,l}$ is the radio channel experienced by the RE at the l^{th} OFDM symbol and the k^{th} subcarrier. The complex exponential $e^{j\phi(k,l,\nu,\tau)}$ results in phase rotations across subcarriers and symbols. The rotations are undesired because they affect the coherency across the OFDM resource grid, but they can be used to estimate the time and frequency offsets in the frequency domain, i.e. after OFDM demodulation. The phase of the exponential is given by:

$$\phi(k,l,\nu,\tau) = 2\pi(lT_{\text{SB}}\nu + kf_\delta\tau) \ . \tag{3.4}$$

To compare the effect of the TO and CFO on the phase term $\phi(k,l,\nu,\tau)$ of (3.4) with the one of the SFO described in [35] and [36], the sampling period $T_s = \frac{1}{f_s}$ is assumed to have an offset ΔT_s. Assuming a common frequency reference, the relative errors on the sampling frequency ξ and the carrier frequency f_c are proportional and relate to the offset of the sampling period:

$$\xi = \frac{\nu}{f_c} \approx \frac{\Delta T_s}{T_s} \ . \tag{3.5}$$

Using the notation introduced in this chapter to rewrite the phase of the complex exponential of Equation (6) in [36] yields:

$$\phi(k,l,\nu,\xi) = 2\pi lT_{\text{SB}}(\nu + kf_\delta\xi) \ . \tag{3.6}$$

The effect of the CFO ν and the SFO ξ on the phase shift across symbols can be compared for the highest subcarrier index $k = N_{\text{FFT}}/2$ using (3.6), $\nu = \xi f_c$ from (3.5), and $f_s = f_\delta N_{\text{FFT}}$:

$$\epsilon_\nu = \frac{\nu}{kf_\delta\xi} = \frac{\xi f_c}{\frac{N_{\text{FFT}}}{2}f_\delta\xi} = \frac{2f_c}{f_s} \ .$$

Assuming values $f_c = 1\,\text{GHz}$ and $f_s = 1.92\,\text{MHz}$ commonly used in NB-IoT and eMTC, the effect of the CFO on the rotation across OFDM symbols is about 1000 times larger than the one of the SFO and can be neglected.

The effect of the TO τ and the SFO on the rotation across subcarriers can be compared similarly with the respective terms in (3.4) and (3.6):

$$\epsilon_\tau = \frac{kf_\delta\tau}{kf_\delta lT_{\text{SB}}\xi} = \frac{\tau}{lT_{\text{SB}}\xi} \ .$$

The accumulated TO due to the SFO after a tracking period of T_T seconds is $\tau = \xi T_T$. Assuming a tracking period of 500 ms, enough to track temperature-driven timing and frequency changes in low mobility scenarios, the effect of the TO on the 10^{th} OFDM symbol is 700 times larger than the one of the SFO and has a negligible impact.

3.3 CRLB for Time and Frequency Offset Estimation

The mathematical model for the effect of synchronization mismatches on OFDM signals can be used to derive the Cramer-Rao lower bound (CRLB) on the variance of the estimated TO and CFO for a given set of pilot signals. The CRLB is a lower bound on the variance of any unbiased estimator of a set of signal parameters and is generally valid for deterministic signals in AWGN [37]. For the considered problem, it gives an insight into the achievable synchronization performance and allows the comparison of different pilot signals according to their expected estimation accuracy.

Assuming a static channel $H_{k,l} = Ae^{j\alpha}$ with amplitude A and phase α, and neglecting the ICI term in (3.3), the signal after matched reception for known modulation symbols with unity power $|X_{k,l}|^2 = 1$ simplifies to:

$$S(k,l;\boldsymbol{\theta}) = X_{k,l}^* Y_{k,l} = Ae^{j(\phi(k,l,\nu,\tau)+\alpha)} + N_{k,l} \ . \tag{3.7}$$

For complex white gaussian noise $N_{k,l}$ with variance σ^2, the SNR can be defined as $\text{SNR} = \frac{A^2}{\sigma^2}$ and the time and frequency offset estimation problem can be handled as the estimation of the unknown signal parameters $\boldsymbol{\theta} = [A, \alpha, \nu, \tau]^{\text{T}}$ in AWGN. In this case, the CRLB is given by the diagonal elements of the inverse of the fisher information matrix $\boldsymbol{J}$ [38]:

$$J_{i,j} = \frac{1}{\sigma^2} \sum_{k,l} \left[\frac{\partial S_I}{\partial \theta_i} \frac{\partial S_I}{\partial \theta_j} + \frac{\partial S_Q}{\partial \theta_i} \frac{\partial S_Q}{\partial \theta_j} \right] \tag{3.8}$$

$S_I = S_I(k,l;\boldsymbol{\theta})$ and $S_Q = S_Q(k,l;\boldsymbol{\theta})$ represent the real and imaginary parts of $S(k,l;\boldsymbol{\theta})$, respectively. The partial derivatives needed to determine the elements of $\boldsymbol{J}$ can be calculated by straightforward

derivation, as shown for the real part S_I in App. A. With the corresponding results for S_Q, the fisher information matrix and its inverse can be calculated numerically for any set of pilots mapped to the OFDM resource grid by setting to zero the terms corresponding to REs that do not carry pilot symbols in the sum of (3.8).

In OFDM systems, pilot signals are commonly distributed in so-called comb-type arrangements where symbols are not mapped to a block of adjacent subcarriers and symbols, but rather distributed to occupy a larger area of the resource grid without filling all REs. This kind of mapping reduces the signalling overhead needed for the characterization of the whole OFDM grid in the time and frequency domains, at the cost of a reduced granularity in the observations that can be made with the pilot signals. In the example mapping depicted in Fig. 3.3, LK pilot signals are evenly distributed every Δ_k subcarriers and Δ_l symbols.

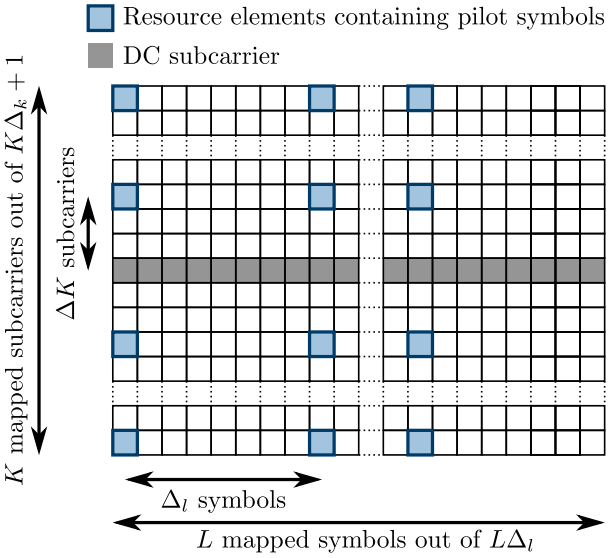

Figure 3.3: Example mapping of pilot symbols to an OFDM grid.

For the regular mapping of Fig. 3.3, the CRLBs on the estimation of time and frequency offsets can be approximated in closed form for relatively large values of K and L as in [38][2]:

$$\mathrm{CRLB}_{\nu}(K,L) \approx \frac{1}{\mathrm{SNR}} \frac{3}{2\pi^2 T_{SB}^2 \Delta_l^2 K L^3} \tag{3.9}$$

$$\mathrm{CRLB}_{\tau}(K,L) \approx \frac{1}{\mathrm{SNR}} \frac{3}{2\pi^2 f_{\delta}^2 \Delta_k^2 L K^3} \ . \tag{3.10}$$

With these closed forms, the impact of signal bandwidth and duration on the CRLBs can be easily identified. The lower bound for the FO in (3.9) decays linearly with the number of occupied subcarriers K in the frequency domain and cubically with the number of used symbols L in the time domain, while for the TO in (3.10) the decay is cubic with the number of subcarriers and linear with the number of symbols. Increasing the subcarrier and symbols spacing between pilot signals quadratically reduces the lower bound of the time and frequency offset estimation, respectively. However, increasing the distance between pilot signals comes at the cost of reduced estimation ranges. In general, for the estimation of the FO longer pilot signals are more performant, while for the TO pilot signals that occupy a larger bandwidth are better. Similar conclusions can be found in literature where time and frequency offset estimation are considered independently [37].

3.3.1 Evaluation of Cellular IoT Pilot Signals

The general formula of (3.8) can be used to evaluate the lower bounds of the residual time and frequency offset errors when estimated with different sets of pilot signals mapped to at least two different OFDM symbols. The resulting RMS error bounds, i.e. the square root of the CRLB, are shown in Fig. 3.4 for synchronization and reference signals available in cellular IoT standards. The plots were generated considering a single subframe and the maximum UE receiver bandwidth for each set of pilots. For NB-IoT, the values for NSSS and NRS are plotted, while for eMTC, the lower bounds for PSS and SSS combined and for CRS are shown. PSS and SSS are considered

[2]Derivations can be found in App. A.2.

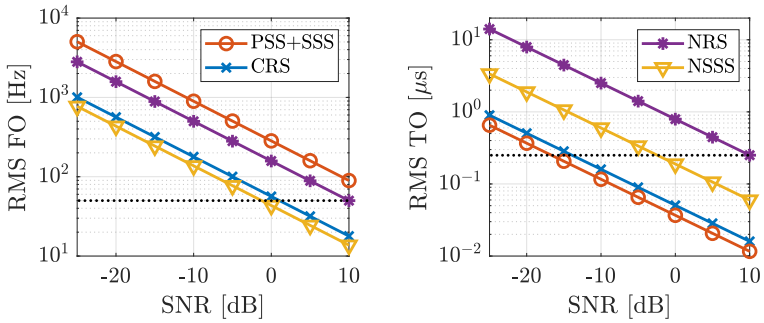

Figure 3.4: Expected lower bound of the RMS error of the frequency (left) and time (right) offset estimation with unknown phase.

together because they can be used jointly since they are mapped to the same antenna port within a subframe. For NB-IoT, NSSS and NPSS can both be used for tracking. The two sequences have similar properties in AWGN. Only NSSS is discussed here because it occupies more subcarriers, offering slightly better performance. Furthermore, NSSS is a better choice in interference-limited scenarios due to its higher resilience to interference from neighbouring cells. As suggested by their lower RMS error (RMSE) on the TO plots on the right, eMTC signals yield better time estimation accuracy than NB-IoT due to their larger bandwidth. On the other hand, NSSS provides a smaller FO RMS error than PSS and SSS in the left plots of Fig. 3.4 due to its longer duration.

To maintain the performance of the radio link and to ensure that the synchronization of the UL signal stays within the limits specified by 3GPP, the FO should be kept lower than 100 Hz and the UE timing should be synchronized within approximately one sample at 1.92 MHz. To reflect these requirements, RMSE lower bounds of 50 Hz and 0.25 µs are targeted in the subsequent analysis for the FO and TO estimation, respectively. These limits are illustrated by the black dotted lines in Fig. 3.4. Since all CRLBs exceed the mentioned values at the target −18.5 dB SNR, signal accumulation is needed. As the bandwidth of the received signal is either imposed by the standard or limited by the UE's receiver, accurate estimation at low SNRs can

only be achieved by accumulating over time. Especially, precise time offset estimation is particularly challenging in NB-IoT and requires very long accumulation periods. This is partially compensated by a relaxed time synchronization requirement for UL transmissions, but cannot be completely avoided in applications that rely on the accuracy of time offset estimation like OTDOA.

To estimate the accumulation time needed to reach the targeted estimation accuracy, a formula for the RMSE CRLB depending on the number of combined subframes is needed. If n_{SF} different SFs are handled as independent observations, the lower bound is inversely proportional to n_{SF} and can be calculated from the value for a single subframe [37] which can be calculated numberically with (3.8):

$$\mathrm{CRLB}_{\nu}(n_{\mathrm{SF}}) = \frac{1}{n_{\mathrm{SF}}}\mathrm{CRLB}_{\nu}(1) \tag{3.11}$$

$$\mathrm{CRLB}_{\tau}(n_{\mathrm{SF}}) = \frac{1}{n_{\mathrm{SF}}}\mathrm{CRLB}_{\tau}(1) \tag{3.12}$$

However, if the n_{SF} subframes are received coherently, the Fisher information matrix must be calculated as formulated in (3.8) for the whole received signal and inverted to determine the CRLBs. To simplify the analysis, the formulas from (3.9) and (3.10) can be used instead when subframes are received consecutively. Even though the mappings used for NRS and CRS are slightly different than the one depicted in Fig. 3.3 and NRS use a rather small number of subcarriers, (3.9) and (3.10) yield valid approximations for these two types of reference signals. However, the formulas do not apply for PSS, SSS, NPSS and NSSS because these pilot signals can be transmitted on different antenna ports across SFs and because they are not located in consecutive SFs, de facto limiting the rate at which the sinusoid in (3.7) is sampled if considered coherently and reducing the FO estimation range.

Fig. 3.5 depicts the number of SFs required to achieve RMSE lower bounds of 50 Hz for the FO and 0.25 µs for the TO. The number of SFs was estimated using (3.9) and (3.10) with $K = 12$, $\Delta_k = 6$, $L = 4n_{\mathrm{SF}}$, $\Delta_l = 3.5$ for CRS, and with $K = 4$, $\Delta_k = 3$, $L = 2n_{\mathrm{SF}}$, $\Delta_l = 7$ for NRS. For the other signals (3.11) and (3.12) were used. The number of SFs needed for time and frequency offset estimation directly impacts the power consumption of the UE during a tracking attempt,

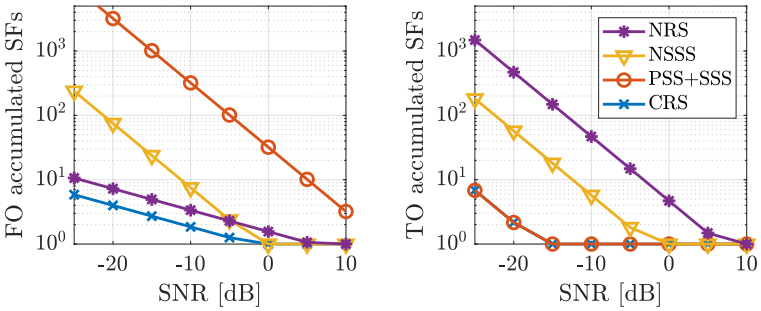

Figure 3.5: Required number of SFs to achieve RMS estimation errors below 50 Hz for the FO (left) and 0.25 µs for the TO (right).

mainly because its receiver has to be switched on for all accumulated SFs. Furthermore, the UE has to stay in an active state during the whole accumulation period and, even if the receiver is switched off, other parts of the system also contribute to the power consumption. Thus, the total acquisition time is also relevant, besides the number of received SFs.

For NB-IoT, fine synchronization with NSSS or NPSS requires tens of SFs for both time and frequency offset detection at low SNRs. As the signals are mapped every 10 ms or 20 ms, the resulting acquisition time is in the order of hundreds of milliseconds. Even though the impact on the acquisition time is compensated by the larger number of SFs carrying NRS, NSSS and NPSS are better suited for time synchronization. As for FO estimation in NB-IoT, NRS are advantageous compared to synchronization signals when more than two SFs are accumulated. This advantage comes from the different behaviour between coherent and non-coherent combination described in (3.9) and (3.11), respectively. This observation is only valid if the accumulated NRS subframes are consecutive. However, not all NB-IoT SFs contain NRS and the mapping depends on the configuration of the cell. Considering the crossing point between NSSS and NRS curves in the FO plot of Fig. 3.5, NRS are better than NSSS for FO estimation if they can be grouped in blocks of three or more consecutive SFs.

For eMTC, TO estimation with CRS and with PSS and SSS combined needs almost the same amount of accumulated SFs, shown by the superimposed red and blue curves in the right plots of Fig. 3.5. The CRS have the advantage of a smaller acquisition time because they are mapped in every SF while PSS and SSS are only mapped every 5 SFs. Regarding FO estimation, CRS outperform PSS and SSS because they can be combined coherently across SFs and exploit the cubic dependence of the CRLB on the accumulation length, providing a considerable advantage in terms of both power consumption and acquisition time at low SNRs.

3.4 Fine Synchronization in eMTC

In the remainder of this chapter, practical solutions for fine time and frequency synchronization in extended coverage are investigated. The presented algorithms are assessed specifically for eMTC with Monte Carlo simulations and their estimation performance is compared with theoretical lower bounds like the ones analyzed in Sec. 3.3. An overview of the different approaches available in the literature for OFDM and LTE systems is given first. Then, the simulation environment used for the evaluations is briefly introduced. The first considered algorithm uses cross-correlations of the received samples with different time- and frequency-shifted replicas of the expected signal for maximum likelihood (ML) estimation with the combination of the PSS and the SSS. As already suggested by the analysis of the CRLB for the different pilots, these synchronization signals provide limited FO estimation accuracy. The framework developed in Sec. 3.2 is used to derive TO and FO estimation techniques using the received symbols after OFDM demodulation. The legacy approach based on phase differences between REs containing CRS proposed for state-of-the-art LTE receivers [39] is extended to support TO estimation. Even though this solution provides sufficient accuracy at low SNRs, it provides diminishing returns for long accumulation periods. Thus, an ML approach for fine synchronization with CRS is presented and combined with complexity reduction techniques to allow its usage on low-cost modem SoCs.

3.4.1 Overview of Tracking Algorithms

As already mentioned in Sec. 3.1, time and frequency offsets of OFDM signals can be estimated in the time domain (TD), on the samples $r_l[n]$ coming from the ADC, or in the frequency domain (FD), on the symbols $Y_{k,l}$ after OFDM demodulation. TD techniques can be used to estimate large time and frequency offsets while FD approaches need preliminary coarse time synchronization for the OFDM demodulation to work well. The supported ranges of FD estimators, limited by subcarrier and symbol spacings, are generally smaller than the ones of TD estimators that are mainly limited by the used sampling rate and the periodicity of the received signal.

Widely used TD tracking techniques, which are also applicable for initial acquisition, use the CP to estimate time and frequency offsets as described in [40]. CP-based TO and FO estimation has the advantages of being a blind estimation technique, which does not need any known pilot signals; supporting large time and frequency offsets; and allowing low-complexity autocorrelation (AC) based implementations. It was also adopted for LTE receivers showing sufficient performance for coarse synchronization at moderate SNR levels [41, 42]. However, the accuracy of *CP AC* estimators drops quickly with decreasing SNR and can only be partially recovered with averaging [33]. Furthermore, the performance of these estimators depends on the number of DL resources used by the BS [14]. To save power, most REs are only used by the BS when they are needed [43]. Thus, the estimation accuracy deteriorates in cells with low data traffic due to the reduced total signal power in the DL.

Another class of TD estimators use the available pilot signals to achieve an ML estimation based on the cross-correlation (CC) of the incoming signal with different time and frequency shifted replicas of a known pilot signal. In [44] an ML method for NB-IoT coarse synchronization using the NPSS is presented. Similar estimators can be implemented for fine synchronization by increasing the density of the tested grid of offsets. For eMTC, it is possible to use the combination of the PSS and SSS to implement a similar *PSS-SSS ML* fine estimation method [13]. Even though the range of searched TOs and FOs is smaller for tracking than for initial acquisition, TD CC-based approaches have a high computational complexity due to

the density of the search grid required to achieve high accuracy. Furthermore, sample-based processing is not efficient for pilot signals that are scattered in a comb-type arrangement in the frequency domain, like CRS for eMTC or NRS for NB-IoT. In this case, it is better to operate on the symbols after OFDM demodulation when possible.

Thus, FD methods are often preferred for tracking if the initial acquisition step provides sufficient synchronization. Even though blind estimation on the OFDM demodulated symbols is possible, e.g. assuming independent and identically distributed (IID) phase-shift keying (PSK) modulated symbols as in [45], pilot-based approaches are more commonly used for cellular communications, where known pilot signals are periodically available [35, 36]. Since in LTE the same pilots are used for channel estimation as for tracking, the authors of [39] and [46] propose estimators based on the phase difference (PD) across received CRS symbols, for which prior channel knowledge is not required. *CRS PD* estimators work well for legacy LTE UEs that do not operate below $-10\,\mathrm{dB}$ SNR but are not well suited for NB-IoT and eMTC UEs that have to support extended coverage because they suffer from noise enhancement and cannot fully exploit accumulation at low SNRs.

To overcome the limitations of legacy CP AC and CRS PD estimators at low SNRs while avoiding the high computational cost of an ML approach, *PBCH* repetitions are used in [33] for fine FO estimation in eMTC. The PBCH estimator is shown to outperform CP AC and CRS PD methods at low SNRs. However, this approach is only applicable when PBCH repetitions are available, and these are not mandatory for the BS. Furthermore, PBCH is only present on the center narrowband. An eMTC UE that supports only the minimum receive bandwidth of $1.4\,\mathrm{MHz}$ may not be able to receive it if it must switch to another narrowband to decode control or data channels.

In this work, the PSS-SSS ML and the CRS PD approaches are evaluated and serve as a baseline for the analysis of the proposed CRS reduced-complexity (RC) ML-based method introduced in the next section.

3.4.2 Simulation Environment

All the considered algorithms were evaluated with a MATLAB© [47] simulation framework including the steps depicted in Fig. 3.6. The digital baseband (DBB) signal of the base station (BS) is generated at $f_s = 1.92\,\text{Msps}$ with a 128-FFT. The digital signal for the l^{th} symbol is:

$$s_l[m] = \frac{1}{\sqrt{N_{\text{FFT}}}} \sum_{k=-\frac{N_{\text{FFT}}}{2}}^{\frac{N_{\text{FFT}}}{2}-1} X_{k,l} e^{j2\pi \frac{k(m-N_{\text{CP}}-lN_{\text{SB}})}{N_{\text{FFT}}}} \quad . \tag{3.13}$$

Compared to (3.1) the CP and symbol lengths were replaced by the corresponding number of samples $N_{\text{CP}} = f_s T_{\text{CP}}$ and $N_{\text{SB}} = f_s T_{\text{SB}}$. Again, $s_l[m] = 0$ for time indices outside of the l^{th} symbol. The overall transmit signal is the sum of all symbols throughout the simulated period:

$$s[m] = \sum_l s_l[m] \ .$$

For multipath fading channels, the signal is first passed through the MATLAB© channel simulator `lteFadingChannel` [48] which supports the three 3GPP LTE channel delay profiles, namely EPA, EVA, and ETU. Otherwise, for static channel simulations, the signal is

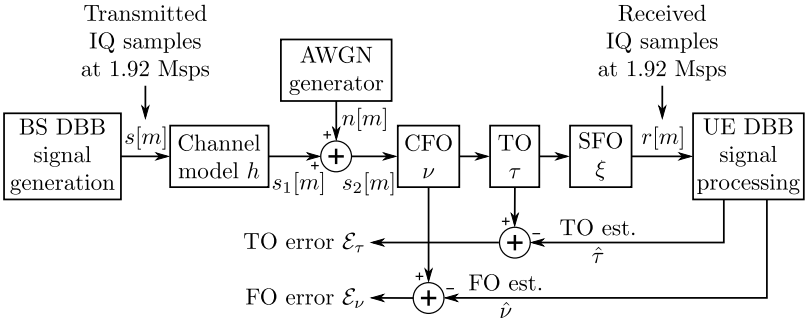

Figure 3.6: Overview of the simulation environment for performance evalutions.

forwarded without modifications, $s_2[m] = s_1[m]$, to the additive noise introduction step:

$$s_2[m] = s_1[m] + n[m]$$

The noise is modelled as a series of complex white Gaussian IID random variables with variance σ^2, i.e. $n[m] \sim \mathcal{CN}(0, \sigma^2)$. The taps of the fading channel are scaled so that they do not alter the expected power of the signal. The SNR is defined as $\frac{1}{\sigma^2}$, which is equal to the transmit signal power divided by the noise power if all the QAM symbols $X_{k,l}$ in (3.13) are modulated with unity average power. As specified, the SNR corresponds to the definition used in 3GPP specifications, the average received power per RE carrying CRS or other unity power symbols divided by the power spectral density of the noise normalized to the subcarrier spacing of 15 kHz [16].

Finally, the effects of the time and frequency offsets are introduced on the noisy signal. Neglecting the drift introduced by the simulated sampling frequency offset (SFO) ξ, the signal received by the UE is:

$$r[m] = s_2[m - \theta] e^{\frac{2\pi\nu(m-\theta)}{f_s}} + n[m - \theta] \ .$$

The sample TO and the carrier frequency offset (CFO) are drawn from uniform distributions in the open intervals $\theta \in]-\theta_{\mathrm{MAX}}, \theta_{\mathrm{MAX}}[$ and $\nu \in]-\nu_{\mathrm{MAX}}, \nu_{\mathrm{MAX}}[$ for selected maximum absolute time and frequency offset values. The SFO is derived from the CFO assuming a given carrier frequency f_c as described in Sec. 3.2:

$$\xi = \frac{\nu}{f_c} \ .$$

The time offset τ in seconds can be calculated using the sample offset with $\tau = \frac{\theta}{f_s}$. The integer part of θ is introduced by cyclically shifting the signal. The edge effects caused by this method do not affect the evaluation of tracking algorithms because the simulated time offsets are smaller than the CP and the last OFDM symbols are not used. For larger offsets, a sufficient margin has to be left unused at the beginning and at the end of the simulated time interval. The fractional part of the TO is achieved by a combination of upsampling and linear interpolation of the signal.

The DBB signal processing algorithms to be implemented on the UE operate on the received signal modelled by $r[m]$. Among others,

the described framework can be used to simulate BLER performance on data channels, initial cell acquisition, or fine synchronization for tracking, which is discussed in this section. The different evaluated algorithms generate estimations $\hat{\tau}$ and $\hat{\nu}$ of the time and frequency offsets, respectively. The estimated values are evaluated by comparing them to the actually introduced offsets. The set of TO and FO estimation errors $\mathcal{E}_\tau$ and $\mathcal{E}_\nu$ obtained after repeated iterations of the simulation process, with different realizations of the involved random variables, are used to generate RMSE and hit or miss rate curves.

3.4.3 ML Synchronization with PSS and SSS

Time and frequency offsets can be estimated with any known signal sequence $s[n]$ of length N_s using a cross-correlation detector on the received samples $r[n] = r(t)|_{t=nT_s}$. The correlation metric for the i^{th} signal period with FO and TO candidates of $\tilde{\nu}$ hertz and $\tilde{\tau} = \tilde{\theta}T_s$ seconds, respectively, is [44]:

$$C_i(\tilde{\nu}, \tilde{\theta}) = \sum_{n=0}^{N_\mathrm{s}-1} r[n+\tilde{\theta}] s^*[n] e^{-j2\pi\tilde{\nu}nT_s} \ .$$

An improved metric is achieved by accumulating the correlations across N_i signal periods non-coherently:

$$C(\tilde{\nu}, \tilde{\theta}) = \sum_{i=0}^{N_i-1} |C_i(\tilde{\nu}, \tilde{\theta})|^2 \ .$$

The joint ML estimator of the time and frequency offsets is then given by the pair of candidates $\tilde{\nu}$ and $\tilde{\theta}$ that yields the highest correlation value:

$$(\hat{\nu}, \hat{\theta}) = \arg\max_{\tilde{\nu}, \tilde{\theta}} C(\tilde{\nu}, \tilde{\theta}) \ . \tag{3.14}$$

In practice, this estimator can be implemented by testing a finite set of candidate pairs. Its computational complexity increases with the number of tested candidates and with the length of the correlation sequence.

The simulated performance of the FO and TO estimator in (3.14) applied to eMTC using the combination of PSS and SSS as the correlation sequence is shown in Fig. 3.7. The RMS FO and TO estimation

errors are plotted in different colours for varying numbers of accumulated subframes N_{SF}, and compared to the corresponding CRLBs represented with dashed lines. Since PSS and SSS are mapped every 5 ms the actual acquisition time in milliseconds is five times higher than N_{SF}. The FO and TO are simulated with maximum values of $\nu_{\mathrm{MAX}} = 1600\,\mathrm{Hz}$ and $\theta_{\mathrm{MAX}} = 3.5$ samples to cover the offsets expected during tracking and to allow a fair comparison with the methods presented later in this chapter that support smaller ranges. The search grid was chosen to cover the mentioned offsets. The TO candidates $\tilde{\theta}$ are equally spaced every sample from -5 to 5 samples. The FO candidates $\tilde{\nu}$ are uniformly distributed every 25 Hz from -1900 to 1900 Hz.

The gain of increasing the number of subframes used for the estimation can be quantitatively evaluated in Fig. 3.7 by comparing the SNRs at which the lines cross a reference RMSE value, curves with better performance being located left.

The accuracy of the TO estimation in the lower plot of Fig. 3.7 never reaches the CRLB, even at high SNRs, because it is limited by the one-sample resolution of the TO search grid. Furthermore, as SNR decreases, there is a threshold after which the RMSE curve quickly deviates from the low error floor and tends to the performance of a random estimator. This threshold effect was studied in [38] for the parameter estimation of a single-tone sinusoid but it applies to ML estimators in general. The threshold effect appears when the RMSE is dominated by outliers in the FO estimation. Nevertheless, the target TO estimation accuracy of $\sim 0.25\,\mu\mathrm{s}$, depicted by the black dotted line, can already be achieved with one SF at -10 dB SNR. Even though each doubling of the number of accumulated signal periods provides less than 3 dB gain, sufficient accuracy can be achieved down to -20 dB SNR when 128 SFs are accumulated.

On the other hand, the FO estimation performance in the upper plot of Fig. 3.7 is not accurate enough for tracking in eMTC, even if it closely follows the CRLB and almost achieves an ideal accumulation gain. The simulated performance is slightly better than the CRLB because the cross-correlation sequence in the time domain includes the CP which is not considered in the CRLB that was calculated on the received signal after OFDM demodulation, i.e. after CP removal. With 128 accumulated subframes the target accuracy of $\sim 50\,\mathrm{Hz}$,

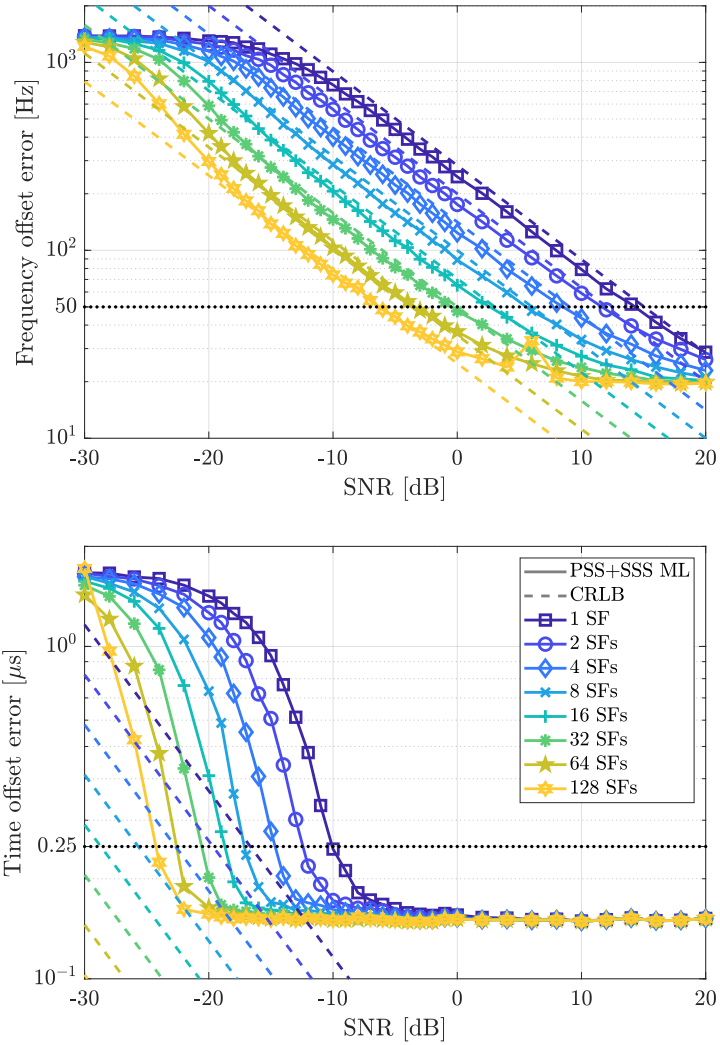

Figure 3.7: FO (top) and TO (bottom) estimation performance of ML PSS and SSS algorithm for different numbers of accumulated subframes compared to corresponding CRLBs in dashed lines. Simulated with 5000 iterations per SNR point.

highlighted with a black dotted line as well, can only be reached down to −6 dB SNR. Assuming the accumulation gain stays constant, at least 4000 SFs would be needed at −20 dB SNR, corresponding to an unaffordable FO acquisition time of 20 s.

The PSS and SSS FO estimation performance also suggests that it is beneficial to relax the accuracy requirement of the initial coarse FO acquisition, using fine estimation algorithms that support larger offsets. As the cell ID-dependent SSS sequence is not known when the initial synchronization is performed, only PSS or blind estimation techniques can be used for this step. In both cases, reducing the residual FO estimation error has a high penalty on the accumulation time, and may even be impossible at low SNR. Taking the PSS-SSS ML method as an example, at −10 dB SNR an accuracy of 100 Hz requires the accumulation of 32 SFs while 200 Hz can already be achieved with 8 SFs.

Overall, as already suggested by the considerations on the CRLB in Sec. 3.3, ML estimation with PSS and SSS can be used for fine TO estimation, but only provide limited accuracy for fine FO tracking. Furthermore, PSS and SSS are only present in the center narrowband and cannot be received by a UE with minimum receive bandwidths, if the UE must receive a physical channel on another narrowband at the same time. CRS-based methods can be used instead, because they provide better SF accumulation gain for the accuracy of the FO estimation, and because CRS are available across the whole cell bandwidth. ML estimation with PSS and SSS can still be used as an intermediate step after the acquisition of the cell ID, if the initial synchronization error is not within the range supported by CRS-based algorithms, or after long sleep periods for the reacquisition of time and frequency synchronization.

3.4.4 CRS Phase Difference Algorithm

Time and frequency offsets can be estimated by considering phase differences between channel estimations on REs containing pilot signals. The least squares (LS) channel estimation $\hat{H}_{k,l}$ can be obtained from the received symbols $Y_{k,l}$ in (3.3) when the modulating symbol $X_{k,l}$ is known [49], as it is the case for REs containing CRS. Neglecting the ICI and noise terms and using the fact that CRS REs have unity

power $|X_{k,l}|^2 = 1$ (i.e. $X_{k,l}^{-1} = X_{k,l}^*$) the LS channel estimation for the l^{th} symbol and the k^{th} subcarrier is:

$$\hat{H}_{k,l} = X_{k,l}^* Y_{k,l} = H_{k,l} e^{2\pi j (l T_{\text{SB}} \nu + k f_\delta \tau)} \ . \tag{3.15}$$

Assuming a static single tap channel, the channel impulse response in the frequency domain simplifies to a constant factor $H_{k,l} = H$. Then, the only variation between the channel estimations on different REs is the rotation due to the time and frequency offsets. In Fig. 3.8 the LS channel estimations on REs carrying CRS are highlighted in blue on a partial OFDM grid of two PRBs and two SFs.

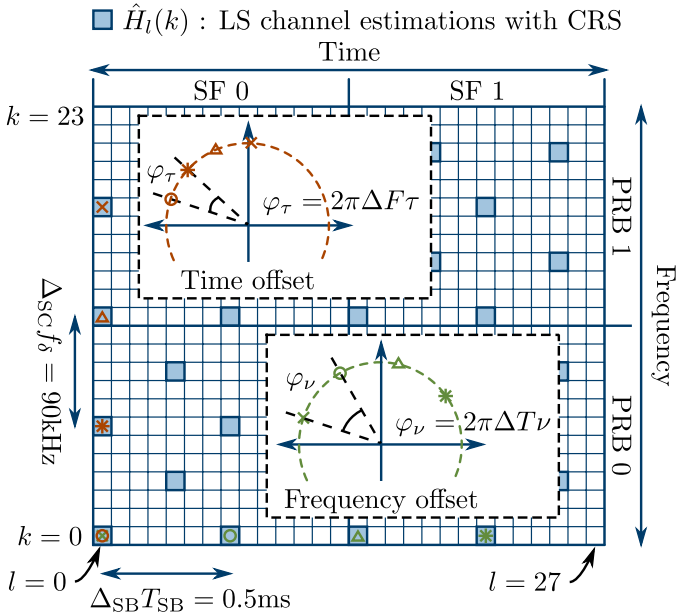

Figure 3.8: Phase drifft across symbols and subcarriers due to frequency and time offsets, respectively.

The phase difference between two LS channel estimations at symbol indices l and λ and subcarrier indices k and κ is composed of two terms, the effects of the FO ν and the TO τ, respectively:

$$\varphi = \mathrm{Arg}\left(\hat{H}^*_{k,l}\hat{H}_{\kappa,\lambda}\right) = 2\pi T_{\mathrm{SB}}\nu(\lambda - l) + 2\pi f_\delta \tau(\kappa - k) \ . \tag{3.16}$$

The varying phase across subcarriers for a fixed symbol index, depicted in the upper inset of Fig. 3.8, is the effect of the TO, while the variation across symbols for a fixed subcarrier, shown in the lower inset, is due to the FO. The two synchronization impairments can be estimated based on the phase difference between neighbour channel estimations with the same symbol or subcarrier index [39, 46]. The phase differences φ_τ and φ_ν can simply be accumulated for all pairs of adjacent CRS to fully exploit the available information because CRS are evenly distributed over frequency and time, every $\Delta_{\mathrm{SC}} = 6$ subcarriers and $\Delta_{\mathrm{SB}} = 7$ symbols, respectively. The resulting time and frequency offset estimations $\hat{\tau}$ and $\hat{\nu}$ are:

$$\hat{\nu} = \frac{1}{2\pi\Delta_{\mathrm{SB}}T_{\mathrm{SB}}}\mathrm{Arg}\left(\sum_{(k,l)\in\mathcal{I}_\nu}\hat{H}^*_{k,l}\hat{H}_{k,l+\Delta_{\mathrm{SB}}}\right) \ ,$$

$$\hat{\tau} = \frac{1}{2\pi\Delta_{\mathrm{SC}}f_\delta}\mathrm{Arg}\left(\sum_{(k,l)\in\mathcal{I}_\tau}\hat{H}^*_{k,l}\hat{H}_{k+\Delta_{\mathrm{SC}},l}\right) \ .$$

The sum for the FO estimation runs over the set of symbol and subcarrier index pairs $\mathcal{I}_\nu$ that carry CRS on all the accumulated SFs, only excluding the last symbols for which $l + \Delta_{\mathrm{SB}}$ would be outside of the considered time interval. The phase between CRS across consecutive SFs is also exploited, assuming that coherency is maintained, i.e. the receiver is not switched off and no frequency hops occur. Similarly, $\mathcal{I}_\tau$ contains all index pairs for which (k, l) and $(k + \Delta_{\mathrm{SC}}, l)$ contain CRS.

The ranges of these estimators are limited by the distance between CRS to $-\hat{\nu}_{\mathrm{MAX}} < \nu < \hat{\nu}_{\mathrm{MAX}}$ and $-\hat{\tau}_{\mathrm{MAX}} < \tau < \hat{\tau}_{\mathrm{MAX}}$ where:

$$\hat{\nu}_{\mathrm{MAX}} = \frac{1}{2\Delta_{\mathrm{SB}}T_{\mathrm{SB}}} = 1\,\mathrm{kHz} \ ,$$

$$\hat{\tau}_{\mathrm{MAX}} = \frac{1}{2\Delta_{\mathrm{SC}}f_\delta} = 5.56\,\mu\mathrm{s} \ .$$

To relax the requirements of the initial synchronization it would be beneficial to also support higher frequency offsets. The time distance between the CRS pairs used to extract the phase change caused by the FO can be reduced by allowing the CRS to be on different but close subcarriers. The frequency range increase with this method comes at the cost of a higher sensitivity of the FO estimator to time offsets.

The RMSE of the legacy CRS tracking algorithm is shown in Fig. 3.9, where the reference RMSE values are highlighter with dotted lines, again. When one or two subframes are used, the FO estimation gets close to the CRLB between 0 and 10 dB SNR, but when the number of repetitions increases, accumulation yields diminishing returns due to the noise enhancement caused by the multiplication used to extract the phase between channel estimations. At low SNRs, the improvement of both TO and FO estimation when the number of accumulated subframes is doubled saturates to 1.5 dB, while an SNR gain of 3 dB per doubling of repetitions is expected for coherent accumulation in AWGN.

Noise enhancement can be explained by taking into account the noise term $N_{k,l}$ of (3.3) contained in $Y_{k,l}$ and inserting it into (3.15). The product of channel estimations in (3.16) is then:

$$\begin{aligned}\hat{H}^*_{k,l}\hat{H}_{\kappa,\lambda} &= \left(X^*_{k,l}Y_{k,l}\right)^* \cdot \left(X^*_{\kappa,\lambda}Y_{\kappa,\lambda}\right) \\ &= |H|^2 e^{2\pi j(T_{\mathrm{SB}}\nu(\lambda-l)+f_\delta\tau(\kappa-k))} \\ &\quad + HN^*_{k,l}e^{2\pi j(\lambda T_{\mathrm{SB}}\nu+\kappa f_\delta\tau)} \\ &\quad + H^*N_{\kappa,\lambda}e^{-2\pi j(lT_{\mathrm{SB}}\nu+kf_\delta\tau)} \\ &\quad + N^*_{k,l}N_{\kappa,\lambda}\ .\end{aligned} \tag{3.17}$$

Assuming a constant channel with unity gain and IID normally distributed complex noise terms with variance σ^2, $N_{k,l} \sim \mathcal{CN}(0,\sigma^2)$, the SNR of the received symbols $Y_{k,l}$ simplifies to $\mathrm{SNR} = \frac{1}{\sigma^2}$ and the variance of the product in (3.17) is:

$$\mathrm{Var}\left[\hat{H}^*_{k,l}\hat{H}_{\kappa,\lambda}\right] = 2\sigma^2 + \sigma^4 \approx \begin{cases} 2\sigma^2, & \text{if SNR} \gg 1 \\ \sigma^4, & \text{if SNR} \ll 1 \end{cases}. \tag{3.18}$$

At low SNR the variance in (3.18) is dominated by the noise product $N^*_{k,l}N_{\kappa,\lambda}$ resulting in a squared noise variance σ^4. The noise enhancement problem arises whenever the multiplication of two noisy terms

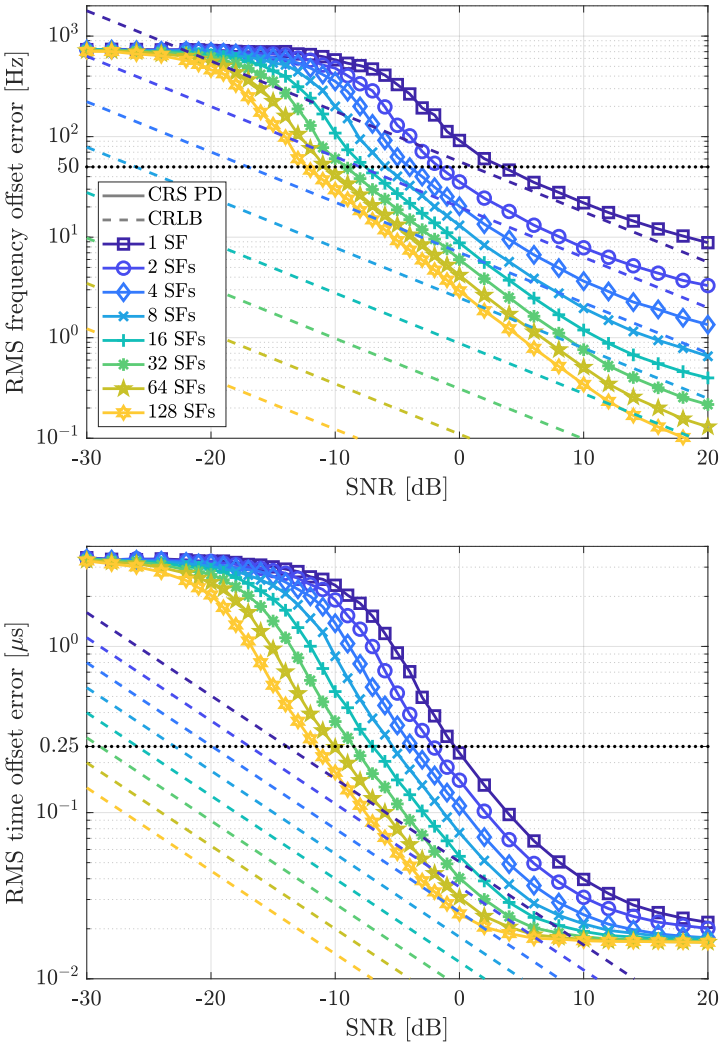

Figure 3.9: FO (top) and TO (bottom) estimation performance of CRS phase difference-based (CRS PD) algorithm for different numbers of accumulated subframes compared to corresponding CRLBs in dashed lines. Simulated with 5000 iterations per SNR point.

is used, which is also the case for CP AC approaches and for the estimator that uses PBCH repetitions proposed in [33]. In terms of SNR, this can be interpreted as negative SNRs being doubled, e.g. if the SNR before the product is $-10\,\mathrm{dB}$, after the product it becomes $-20\,\mathrm{dB}$. In Fig. 3.9 this can be seen from the slopes of the RMSE curves for 128 SFs, both TO and FO estimation errors decay by one decade for a 10 dB increase in SNR, and not 20 dB like the CRLB, or from the gain saturating to 1.5 dB instead of 3 dB when the number of SFs is doubled. To overcome this limitation an ML approach has to be considered also for CRS-based tracking.

3.5 ML-Based Estimation with CRS

Two approaches with different limitations were discussed so far. The performance of the ML-PSS-SSS method for fine synchronization is restricted by the used pilots and the corresponding CRLB on the FO estimation, while the legacy CRS algorithm fails at exploiting accumulation due to noise enhancement. The solution discussed below aims at taking the best of both worlds with an ML estimation using CRS.

3.5.1 2D-FFT Algorithm

The sequence of LS channel estimations of (3.15) can be considered as a two-dimensional complex sinusoid across subcarriers and OFDM symbols. To highlight this behaviour the real parts of the channel estimations are depicted on the OFDM resource grid in the left-hand part of Fig. 3.10, where the REs that do not contain CRS were filled with linearly interpolated values. Thus, the FO and TO estimation can be simplified to a sinusoidal parameter estimation problem. It is well known that the frequency of a single-tone signal can be estimated considering the bin with the highest power of its DFT [38]. In this case, the time and frequency offsets can be estimated by finding the peak of the two-dimensional PSD of the grid of channel estimations.

The PSD can be approximated by the power of a 2D DFT across subcarriers and symbols, illustrated in the right part of Fig. 3.10. There, a 2D DFT of size 512 in both directions was taken on the

grid of LS channel estimations on CRS for 8 SFs and the eMTC UE bandwidth of one narrowband. The REs that do not contain CRS and that are outside of the considered grid were padded with zeros to fill the input of the DFT. The regular peaks that appear in the approximation of the PSD are aliases introduced by the zero padding between CRS locations. The supported TO and FO ranges of an algorithm based on this metric are limited by the output ranges of the 2D DFT or by the distance between the actual peak and the closest secondary peaks. The resolution of the estimated values depends on the FFT size and the distance between the grid elements at the input of the FFT. If the REs containing CRS are equally spaced in one direction, the input of the DFT can be packed along this dimension reducing the required zero padding and DFT size without impacting the performance. For instance, the input can be packed along the frequency axis because CRS are present every 3$^{\text{rd}}$ subcarrier, but not along the time axis because CRS are alternatively mapped every 3 and 4 symbols.

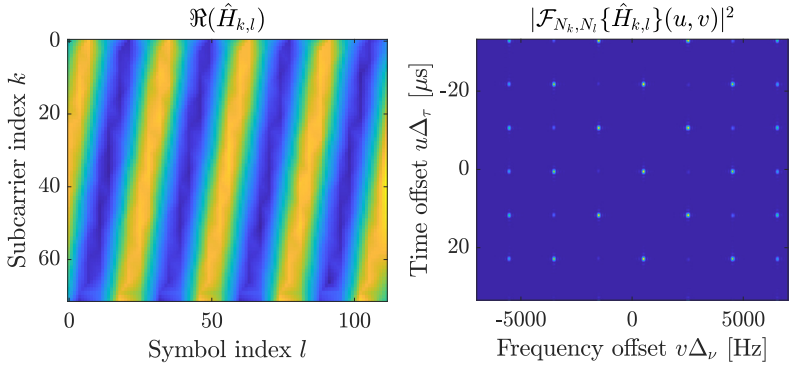

Figure 3.10: Effect of an FO of 500 Hz and a TO of 0.5 µs on the 2D grid of LS channel estimation $H_{k,l}$. The real part of the interpolated values of $H_{k,l}$ is shown left and the power of the 2D 512-DFT across symbols and subcarriers right.

Using the symbol $\mathcal{F}_{N_k,N_l}\{\cdot\}(u,v)$ to represent the N_k by N_l 2D DFT of a zero-padded grid of channel estimations, with discrete output frequencies $u \in \{-\frac{N_k}{2}, ..., \frac{N_k}{2} - 1\}$ and $v \in \{-\frac{N_l}{2}, ..., \frac{N_l}{2} - 1\}$, the maximum search can be expressed as:

$$(\hat{u}, \hat{v}) = \arg\max_{(u,v)} \left| \mathcal{F}_{N_k,N_l}\{\hat{H}_{k,l}\}(u,v) \right|^2 .$$

The FO and TO estimations in hertz and seconds are then:

$$\hat{\nu} = \frac{\hat{v}}{N_l T_{\mathrm{SB}} \Delta_l} \tag{3.19}$$

$$\hat{\tau} = \frac{\hat{u}}{N_k f_\delta \Delta_k} . \tag{3.20}$$

In this case, Δ_l and Δ_k are the numbers of symbols and subcarriers that separate the grid elements of the DFT input along time and frequency, respectively.

This ML estimation algorithm was simulated with DFT dimensions $N_l = N_k = 2048$ and grid spacings $\Delta_l = 1$ and $\Delta_k = 3$. The mentioned parameters result in FO and TO estimation resolutions of $\Delta_\nu = 6.8\,\mathrm{Hz}$ and $\Delta_\tau = 0.011\,\mu\mathrm{s}$, respectively, enough for the targeted accuracy without the need for further interpolation. The results for different numbers of accumulated SFs are shown in Fig. 3.11. As opposed to the legacy PD algorithm the curves follow now the CRLB for both TO and FO, even for negative SNR. Three different operating conditions can be identified from the FO RMSE plots. At sufficiently high SNRs, the performance is limited by the resolution provided by the DFT dimensions and the RMSE saturates to a constant $\frac{\sqrt{3}}{6}\Delta_\nu$. In the intermediate SNR range, the estimation error follows the CRLB. As SNR decreases, there is a point after which the RMSE curve quickly deviates from the CRLB due to the same threshold effect described in Sec. 3.4.3 for the PSS-SSS ML estimator, i.e. because errors are dominated by outliers in the FO estimation [38].

The SNR at which performance collapses is improved by 2 to 3 dB when the number of received SFs is doubled, e.g. on the FO plot, the cut-off SNR level is improved from $-10\,\mathrm{dB}$ with 8 SFs to $-13\,\mathrm{dB}$ with 16 SFs. Furthermore, the FO estimation exploits the advantage of coherent accumulation that was expected from the analysis of the

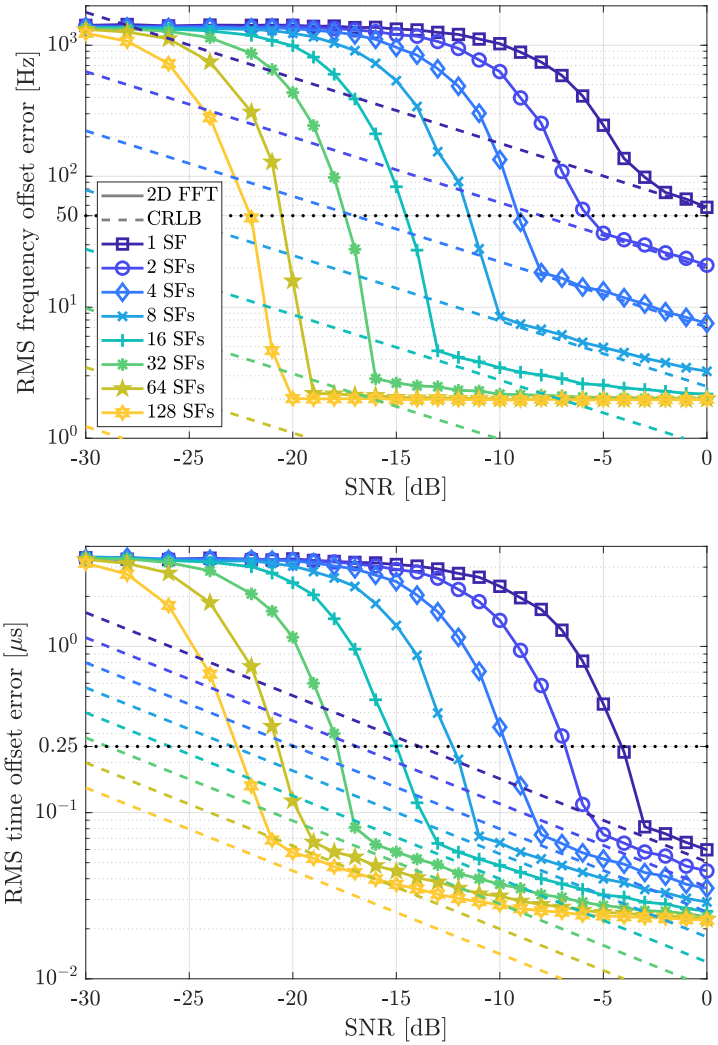

Figure 3.11: FO (top) and TO (bottom) estimation performance of 2D FFT algorithm for different number of accumulated subframes compared to corresponding CRLBs in dashed lines. Simulated with 5000 iterations per SNR point.

CRLB, providing a significant improvement to the estimation accuracy. The accumulation gain starts to saturate between 64 and 128 SFs where the FFT size starts to be a limiting factor. As described in [38], the FFT must be at least twice as long as the considered signal to avoid excessive misdetections of the PSD peak at low SNRs. In this case, a 2048-FFT can contain up to 146 SFs with 14 symbols each, limiting the capture range without accumulation gain degradation due to the FFT size to 73 SFs.

3.5.2 Reduced Complexity ML Algorithm

Even though the 2D DFT can be efficiently implemented using FFTs, the ML algorithm of Sec. 3.5 fully exploits the potential of CRS for TO and FO estimation at very high computational and memory costs. For example, to accumulate 128 SFs the DFT size across symbols should be at least $N_l = 2048$, considering only powers of 2 for FFT sizes for implementation purposes. To cover all subcarriers containing CRS, 24 FFTs of this size are needed. Then, FFTs across subcarriers of minimum size $N_k = 64$ are needed to provide sub-sample resolution. To cover a search range of $\pm 1800\,\mathrm{Hz}$, 526 FFTs across subcarriers are needed. Without considering the power calculation and the maximum search, assuming $5N\log_2(N)$ real operations to compute an FFT of size N, and 4 bytes (B) per stored complex value, this requires $\sim 3.7\,\mathrm{MOP}$ and at least $147\,\mathrm{kB}$ of memory.

To reduce the number of operations and the storage requirement, channel estimations can be combined as depicted in Fig. 3.12. Channel estimations corresponding to CRS located in the same box are added before estimating the offsets on the accumulated values. The underlying assumption is that the time and frequency offsets are small enough not to alter the coherency of the combined channel estimations. This simplification reduces the number of FFTs that have to be computed and the amount of data that has to be stored. However, the supported estimation ranges are also reduced because performance deteriorates for large offsets when the accumulation across subcarriers or symbols is damaged by the TO or the FO, respectively. The larger the section covered by the accumulation box, the smaller the supported estimation ranges. Denoting the set of subcarrier and

symbol indices of the combined channel estimations by $\mathcal{C}_{m,n}$, the m^{th} and n^{th} DFT input in frequency and time domain is:

$$\tilde{H}_{m,n} = \sum_{(k,l)\in\mathcal{C}_{m,n}} \hat{H}_{k,l} \ . \tag{3.21}$$

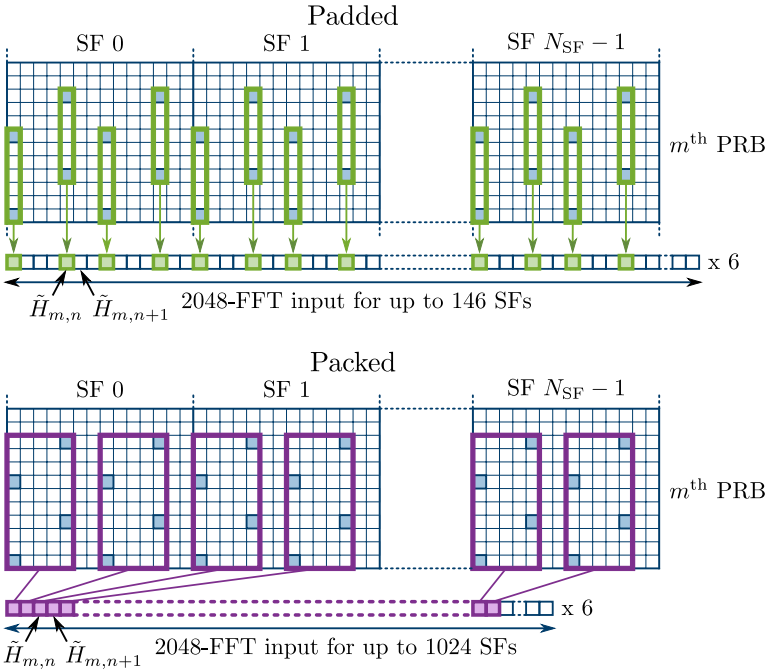

Figure 3.12: Complexity reduction by combination of CRS channel estimations. In the padded version, each time domain FFT input bin corresponds to one OFDM symbol, in the packed version it corresponds to one slot. The mapping is repeated of 6 PRBs.

Two options are considered in Fig. 3.12. The first, denoted as padded and shown in green at the top of the figure, only combines channel estimations that are on the same symbol and the same PRB by groups of two. It only needs $M = 6$ FFTs across symbols and divides the required memory to store the DFT input by two. The FO

estimation range is not affected by the simplifications compared to the 2D-FFT algorithm, because channel estimations are only combined across subcarriers. The supported frequency range is about $\pm 1.7\,\mathrm{kHz}$, limited by the larger of the two regular spacings between consecutive OFDM symbols carrying pilots, i.e. 4 symbols for CRS. On the other hand, the supported TO range is divided by two compared to the 2D DFT approach to approximately $\pm 2.5\,\mu\mathrm{s}$. In the packed alternative, depicted in violet at the bottom of Fig. 3.12, groups of four channel estimations within the same slot and PRB are combined. The memory requirement is divided by two compared to the first option for the same number of accumulated SFs. If the FFT size across symbols is kept constant this solution also allows longer accumulation times of up to 1024 SFs instead of 146. The TO estimation range is the same as for the solution in green, but the FO estimation range is reduced to $\pm 1000\,\mathrm{Hz}$ due to the increased distance between the DFT input bins in the time domain.

Another approach to further reduce the computational complexity is to perform the search in two steps, finding first the FO and then the TO. Like this, the full 2D DFT grid does not have to be computed. Six FFTs are first performed across symbols, but then only one FFT on the bin with maximum power accumulated over the first six FFTs is needed across subcarriers. Representing the one-dimensional DFT of size N_n of a signal by $\mathcal{F}_{N_n}\{\cdot\}(v)$ the two-step algorithm with combined grid elements is:

$$\begin{aligned} \tilde{H}_m(v) &= \mathcal{F}_{N_n}\{\tilde{H}_{m,n}\} \ , \\ \hat{v} &= \arg\max_{v} \sum_{m=0}^{M-1} |\tilde{H}_m(v)|^2 \ , \\ \tilde{H}(u,\hat{v}) &= \mathcal{F}_{N_m}\{\tilde{H}_m(\hat{v})\} \ , \\ \hat{u} &= \arg\max_{u} |\tilde{H}(u,\hat{v})|^2 \ . \end{aligned} \tag{3.22}$$

To adapt this algorithm for multiple antenna ports (APs), the DFTs must be calculated for the channel estimations on each AP and all the DFT powers added before the maximum searches.

The performance of the described two-step approach was evaluated for a single antenna port and the two combination strategies depicted

in Fig. 3.12. The same FFT size $N_{\mathrm{FFT}} = N_m = N_n = 2048$ was used across time and frequency domains. It was chosen to match the size of the FFT hardware accelerator used for initial synchronization described in [13], allowing the reuse of the same accelerator for tracking. The performances of the packed and padded versions for a static AWGN channel are shown in Fig. 3.13 and compared with the results of the 2D-FFT algorithm. The estimation miss rate with 64 accumulated SFs is shown for the three options, while the performance for the packed version of the reduced-complexity (RC) ML CRS estimation is also plotted for 512 SFs because it can accommodate more subframes with the same FFT size. For all FO and TO miss rate plots throughout this chapter, an estimation is counted as a miss if the absolute FO estimation error is larger than 100 Hz or the absolute TO estimation error is larger than one sample, i.e. 0.52 µs. The simulated TO was uniformly distributed with a maximum absolute value of $\theta_{\mathrm{MAX}} = 3.5$ samples. For the packed RC ML algorithm, the maximum absolute FO

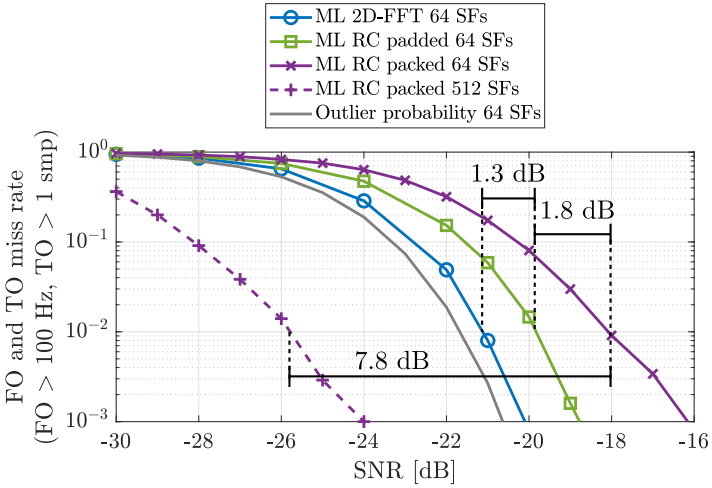

Figure 3.13: Comparison of FO and TO miss rates of different ML-based estimation algorithms. Simulated with 5000 iterations per SNR point.

was $\nu_{\mathrm{MAX}} = 800\,\mathrm{Hz}$ due to its reduced estimation range, while for the other methods a maximum value of $\nu_{\mathrm{MAX}} = 1600\,\mathrm{Hz}$ was used. Due to the high penalty of a wrong correction following a misdetection, the time and frequency offset estimation should be operated at almost zero error rates if no condition for hit detection is used. However, the results are evaluated at 1% miss rates for which sufficient sample sizes can be reached with acceptable simulation time.

The miss rate of the ML estimator follows closely its theoretical outlier probability plotted with a grey solid line. The outlier probability was calculated as described in [38], replacing the length of the signal N with the number of REs carrying CRS in the observed period. The results for the two RC ML methods are slightly worse than for the 2D-FFT approach, with 1.3 dB and 3.1 dB performance loss at a 1% miss rate for the padded and the packed version, respectively. The non-coherent combination of FFT power across PRBs for the FO estimation slightly reduces the accumulation gain compared to the 2D-FFT method, where the coherent accumulation of all PRBs is ensured by the DFTs in the frequency domain. Additionally, the two RC methods suffer from degradation due to the TO, because REs on different subcarriers are directly added in (3.21) even though their phases are different due to the TO. For the packed approach, the performance is also affected by the FO, due to the accumulation of REs located on different OFDM symbols. On the other hand, the packed RC ML algorithm allows longer accumulation times, achieving a 1% miss rate down to -25.8 dB SNR with 512 SFs. This performance corresponds to an improvement of 7.8 dB compared to the accumulation of 64 SFs out of the 9 dB that would be expected from an ideal accumulation gain.

3.5.3 Non-Coherent Accumulation

One limitation of the discussed CRS ML estimators is that they require the accumulated SFs to be coherent, meaning that the UE must continuously receive on the same frequency during the whole observation period. This can be ensured during the fine correction step following initial synchronization or for reacquisition after wake-up from a sleep period. However, in other cases, the coherent reception

could need to be interrupted for uplink transmissions or frequency jumps.

The two-step RC ML algorithms can be adapted to deal with P non-coherent blocks of data of length N by applying separate FFTs to each signal segment. The FO is then estimated by searching for the bin with maximum power accumulated across all FFTs, while the TO is given by the maximum of the accumulated power of P FFTs across $M = 6$ PRBs. The procedure can be summarized in mathematical notation as in (3.22):

$$\begin{aligned} \tilde{H}_m^{(p)}(v) &= \mathcal{F}_{N_n}\{\tilde{H}_{m,pN+n}\} \ , \\ \hat{v} &= \arg\max_v \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} |\tilde{H}_m^{(p)}(v)|^2 \ , \\ \tilde{H}^{(p)}(u,\hat{v}) &= \mathcal{F}_{N_m}\{\tilde{H}_m^{(p)}(\hat{v})\} \ , \\ \hat{u} &= \arg\max_u \sum_{p=0}^{P-1} |\tilde{H}^{(p)}(u,\hat{v})|^2 \ . \end{aligned} \tag{3.23}$$

The index n runs from 0 to $N-1$ to only include the elements located within the same coherent burst in each FFT across the time domain.

The FO and TO estimation miss rates for the padded RC ML estimator with 64 received subframes and different numbers of non-coherently accumulated signal blocks are depicted by the solid lines of Fig. 3.14. The best results are achieved when all SFs are accumulated coherently in one block. The performance loss when halving the number of coherently combined SFs increases for small coherent accumulation times.

The smaller gain of non-coherent combination can be further investigated by considering the impact of coherent and non-coherent accumulation on the outlier probability of the FO estimator of (3.23) in AWGN, following a similar analysis as in [38]. Assuming that the combination of channel estimations in (3.21) is lossless, that no zero padding is needed at the input of the DFT, that a DFT of the same size as the input vector is used, and that the FO corresponds exactly to the frequency of one DFT output bin, the estimation metric for $\hat{u}$ can be shown to have a chi-squared distribution with $\kappa = 2PM$ degrees of freedom on the noise bins, and a non-central chi-squared

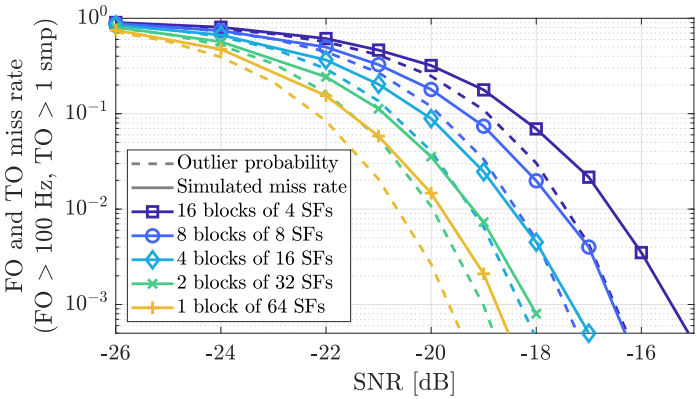

Figure 3.14: Comparison of combined FO and TO miss rates for the CRS ML RC padded estimator with 64 received SFs separated into different numbers of non-coherently accumulated blocks P. Simulated with 10000 iterations per SNR point.

distribution with the same degrees of freedom and a non-centrality parameter $\lambda = \frac{2N_c PM}{\sigma^2}$ on the signal bin, where σ^2 is the noise power, and $N_c = 8N_{SF,coh}$ the number of coherent REs included in each DFT input. The outlier probability can then be obtained by replacing the probability density functions (PDFs) of Equation (57) in [38] with the ones of the distributions mentioned here. The resulting outlier probabilities are plotted in dashed lines in Fig. 3.14 for the same numbers of coherently accumulated SFs as the miss rate plots. The PDFs and the final integral form of the outlier probability are discussed in App. A.3.

The differences between the outlier probabilities and the corresponding FO and TO estimation miss rate plots are due to the loss caused by the TO in the accumulation of (3.21). Considering the relative difference between the curves when $N_{SF,coh}$ is doubled, the dashed and solid lines have similar behaviour, for example, increasing the number of coherently accumulated subframes from 8 to 16 yields an SNR improvement of 0.86 dB for an outlier probability of 1% and 0.84 dB for the corresponding miss rate.

The performance difference between coherent and non-coherent combinations can be understood by comparing the underlying mechanisms with which the two approaches improve the outlier probability with accumulation. Considering the probability distributions of the simplified model used to calculate the outlier probabilities, the expected value of the signal bin E_s and the expected value of the noise bins E_n are:

$$E_s = 2PM\left(1 + \frac{N_\mathrm{c}}{\sigma^2}\right) ,$$
$$E_n = 2PM .$$

Their ratio $\frac{E_s}{E_n} = 1 + \frac{N_\mathrm{c}}{\sigma^2}$ increases with N_c, the number of coherently accumulated REs, but is independent of the number of non-coherent accumulated blocks P. In other words, the ratio between the expected value of the desired peak and the expected value of the noise bins is improved with coherent accumulation, therefore increasing the probability of selecting the correct peak. However, the ratio is not affected by non-coherent accumulation.

To understand how non-coherent accumulation can also help reduce the outlier probability, the variance of the estimation metric relative to the expectation of the peak value squared has to be taken into account. Denoting the variance of the signal bin by V_s and the variance of the noise bins by V_n, their ratios to the expected peak value squared are:

$$\frac{V_s}{E_s^2} = \frac{1}{P}\frac{1 + \frac{2N_\mathrm{c}}{\sigma^2}}{\left(1 + \frac{N_\mathrm{c}}{\sigma^2}\right)^2} ,$$
$$\frac{V_n}{E_s^2} = \frac{1}{P}\frac{1}{\left(1 + \frac{N_\mathrm{c}}{\sigma^2}\right)^2} .$$

In this case, increasing the number of non-coherently combined blocks P reduces the variance of the estimation metric compared to the expected peak value, also reducing the outlier probability, but to a smaller extent than coherent accumulation, as shown by Fig. 3.14.

Even though coherent accumulation yields better results under static channel conditions, non-coherent accumulation combined with frequency hopping could be used to exploit the diversity of frequency

selective fading channels. This could potentially lead to a different balance in the performance gains of coherent and non-coherent accumulation, depending on the number of available independent channel realizations throughout the cell's bandwidth.

3.5.4 Hit Detection

The supported SNR ranges of all the approaches discussed in this chapter were extended by accumulating signals over longer periods. The required accumulation times are highly dependent on the SNR of operation. Using excessive RX times in good channel conditions unnecessarily increases the power consumption of the device. On the other hand, if the UE corrects wrongly estimated errors it may lose connection and need to trigger a connection reestablishment procedure at high power and latency costs. Thus, the tracking algorithms must be operated with sufficient accumulation margin to avoid large estimation errors. The desired trade-off for the accumulation time can be set based on an estimation of the SNR. However, SNR estimation may be tedious in extended coverage and a hit condition that can be easily integrated into the algorithms' flow is preferable. For the RC ML algorithms, the peak-to-average and peak-to-second-peak of the metric used for FO estimation provide sufficient information on the reliability of the detected peak. The mentioned metric was introduced in (3.23) of Sec. 3.5.3 as part of the maximum search for FO estimation:

$$C_\nu(v) = \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} |\tilde{H}_m^{(p)}(v)|^2 \ . \tag{3.24}$$

The estimated FO $\hat{\nu}$ is given by the discrete frequency $\hat{v}$ with the highest metric and the peak, second peak and average values that can be used to calculate the ratios for hit detection can be derived from the same metric for a symmetrical FO search range of size N_v:

$$\begin{aligned} C_{\max} &= C_\nu(\hat{v}) \ , \\ C_{\mathrm{snd}} &= \max_{|\hat{v}-v|>\delta_v} C_\nu(v) \ , \end{aligned}$$

$$C_{\text{avg}} = \frac{1}{N_v} \sum_{v=-\frac{N_v}{2}}^{\frac{N_v}{2}-1} C_\nu(v) \ .$$

For the second peak search, an exclusion range of size $\delta_v = \frac{N_{\text{FFT}}}{N_{\text{bins}}}$ around the maximum is needed to compensate for the width of the main peak caused by the difference between the used FFT size and the number of occupied FFT input bins N_{bins}. The peak-to-average and peak-to-second-peak ratios obtained from these three values can be compared with predefined thresholds to generate hit conditions:

$$\frac{C_{\max}}{C_{\text{avg}}} > \eta_1 \text{ and}$$
$$\frac{C_{\max}}{C_{\text{snd}}} > \eta_2 \ .$$

The ideal values of the thresholds vary depending on which of the RC ML algorithms is used and on the number of antenna ports, subframes, and non-coherent blocks that are considered. In this analysis, the two conditions were combined and the same thresholds were used for the padded and packed versions and all numbers of accumulated SFs from 16 to 1024, considering only coherent accumulation in time, i.e. $P = 1$ coherent block, and a single antenna port. For smaller numbers of SFs higher thresholds are needed due to the smoother and broader peaks of the resulting metrics. The thresholds were set such that no false positive hits were observed in SNR regions with miss rates smaller than 10%. The same thresholds $\eta_1 = 4$ and $\eta_2 = 1.74$ were used for the RF measurements presented in Chap. 6. The simulated miss rates obtained with these thresholds for the padded RC ML method with 64 SFs and the packed RC ML method with 512 SFs are depicted in Fig. 3.15.

The plots generated including the detection thresholds are compared with miss rate curves based only on the estimated FO and TO without hit conditions. Again, 100 Hz and one sample serve as error limits for hits. If the thresholds are not met when a hit condition is used, the FO and TO correction steps can be skipped, and a new estimation attempt can be started or the number of accumulated subframes increased. Therefore, the algorithms can be operated at

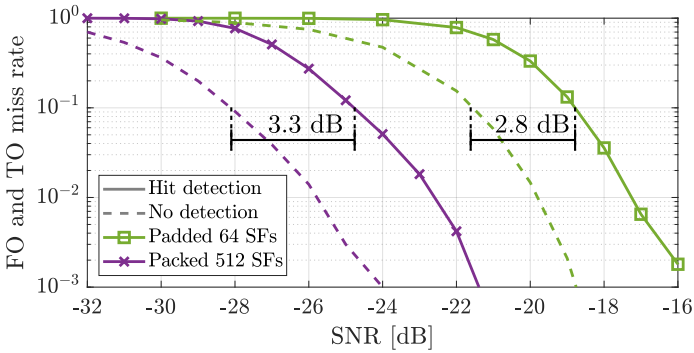

Figure 3.15: FO and TO miss rate using thresholds for hit detection against miss rate based only on the estimated offsets. Simulated with a static AWGN channel and with 10000 iterations per SNR point.

a 10% miss rate. At this error level, the losses of 3.3 dB and 2.8 dB compared to the true miss rates are highlighted in Fig. 3.15 for the padded and packed ML RC algorithm, respectively. The hit detection thresholds are also suited for early termination.

3.6 Performance Evaluation

To conclude this chapter, the overall performance of the proposed ML-based reduced complexity (ML RC) algorithms for fine time and frequency offset estimation will be evaluated and compared with the legacy solutions discussed in previous sections.

3.6.1 Comparison of Algorithms

The miss rates of the different approaches using CRS or a combination of PSS and SSS discussed so far are shown in Fig. 3.16. The solid lines correspond to algorithms evaluated with 64 accumulated SFs and the dashed lines with 512. All methods were evaluated with a single antenna port (AP). The simulated TO was uniformly distributed with a maximum absolute value of $\theta_{\mathrm{MAX}} = 3.5$ samples. For the CRS ML RC packed and the legacy CRS phase difference (PD) algorithms the maximum absolute FO was $\nu_{\mathrm{MAX}} = 800\,\mathrm{Hz}$ due to their reduced estimation ranges. For the other CRS ML RC padded and PSS-SSS ML algorithms a maximum value of $\nu_{\mathrm{MAX}} = 1600\,\mathrm{Hz}$ was used. The same distributions of the time and frequency offsets were used for all the ramaining plots of this chapter.

The performance of the ML RC algorithms includes the hit detection step discussed in Sec. 3.5.4 and was evaluated for coherent accumulation in the time domain which performs better than non-coherent accumulation in static conditions. For the other algorithms, the miss rate was evaluated based only on the estimated offsets. The thresholds for FO and TO misses are 100 Hz and one sample, respectively.

The results for the ML estimation with PSS and SSS show the limits of these pilots for fine synchronization. The performance of this approach is limited by the FO estimation due to the short duration of the considered signal, even when PSS and SSS are combined. On the other hand, the results of Sec. 3.4.3 showed that accurate time synchronization is possible with PSS and SSS at acceptable accumulation times even in extended coverage. Results with PSS alone would lead to similar conclusions. As discussed in the CRLB considerations, the performance difference between time and frequency offset estimation would even be amplified, because the PSS alone has the same bandwidth but reduced length compared to the combined PSS-SSS

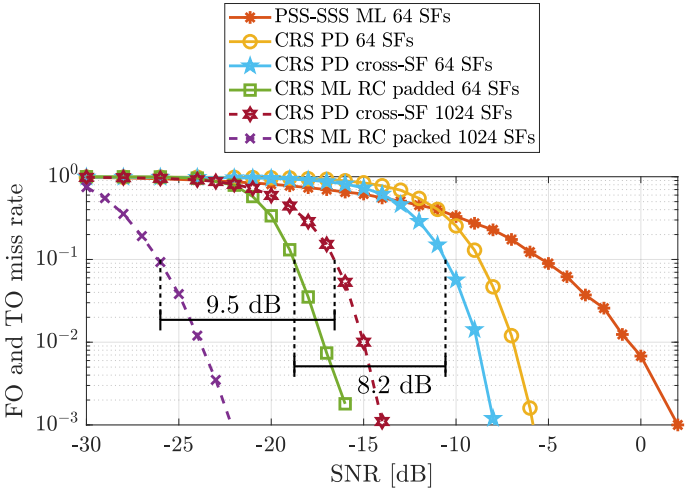

Figure 3.16: Comparison of FO and TO miss rate for different estimation algorithms. Simulated with a static AWGN channel and with 5000 iterations per SNR point.

sequence. Thus, frequency synchronization is harder to achieve than time synchronization during the initial cell acquisition which relies only on the PSS. As mentioned in Sec. 3.4.3, the time needed for the initial synchronization can be reduced by relaxing the requirements on the FO error after the initial acquisition. Higher residual frequency offsets can be supported by increasing the supported frequency range of the tracking algorithm. Therefore, the need to support high FO estimation ranges was taken into account in the development of the CRS ML RC algorithms.

Two versions of the CRS PD algorithm are shown in Fig. 3.16. The first is a straightforward extension of the algorithms available in the literature to support coverage extension of the accumulated complex phases measured in each subframe [39, 46]. This version was also implemented on the SoC and used as a reference point for the RF performance measurements of Chap. 6. The second, CRS PD cross-SF, yields a 1.9 dB gain by also considering the phase differences

of reference symbols across subframes. The best of the two methods is plotted for 64 and 1024 accumulated subframes. Even though CRS PD cross-SF does not include hit detection, it loses 8.2 dB and 9.5 dB at a 10% miss rate compared to the padded and packed RC ML algorithms with the same number of accumulated subframes. The better performance of ML RC algorithms also translates into a reduced acquisition time at the same SNR level as shown in Fig. 3.17.

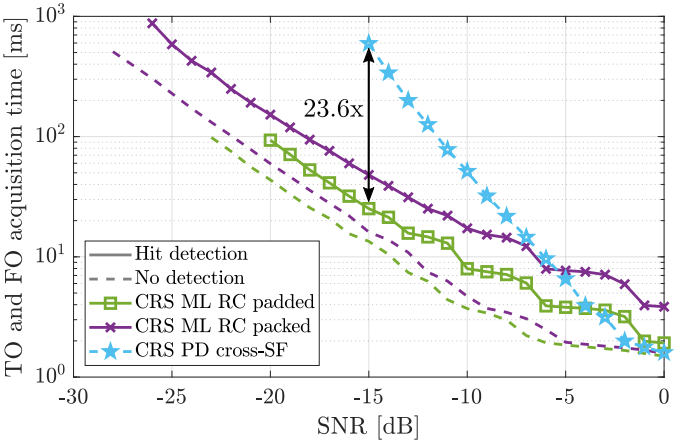

Figure 3.17: Comparison of FO and TO acquisition time to achieve a 10% miss rate for different estimation algorithms with CRS. Simulated with a static AWGN channel and 10000 iterations per SNR point.

The figure shows the accumulation time required to achieve a miss rate of 10% in FO and TO estimation with CRS. The curve was generated by verifying the hit condition after different numbers of received subframes and interpolating between the numbers of SFs that yield the error rates closest to 10%. The ML RC algorithms were evaluated with and without hit detection based on peak-to-average and peak-to-second-peak ratios, while the PD cross-SF algorithm was evaluated with the miss rate of the estimated offsets. The hit detection condition can be used for the early termination of the FO and TO estimation. At −15 dB SNR, the PD algorithm needs 595 ms to reach

a miss rate of 10% while the padded version of the ML RC approach only needs 25.2 ms, 23.6 times less. The difference even increases to 44 times if the ML RC algorithm is considered without hit detection, showing how algorithmic improvements contribute to reducing the power consumption. As already shown in Sec. 3.5.2, the padded ML RC approach yields better performance at the same accumulation level and has a higher FO estimation range compared to the packed one. However, packed ML RC estimation allows longer accumulation times bringing the covered SNR down to −26 dB.

3.6.2 Performance in Fading Channels

So far, only static AWGN channels have been considered because they yield representative results and allow simpler theoretical analysis than more sophisticated channel models. However, radio propagation conditions are often better modelled by multipath fading channels. As already mentioned in Chap. 2, the 3GPP defined three channel models for the performance evaluation of LTE devices and the same are used for NB-IoT and eMTC. In Fig. 3.18, the separated time and frequency offset estimation miss rates of the ML RC padded and the PD cross-SF estimators, both without hit detection, are plotted for the EPA and ETU channels, the ones with lowest and highest delay spread, respectively. The propagation conditions represent low-mobility scenarios with Doppler frequencies of 5 Hz for the EPA channel and 1 Hz for the ETU case.

For a single antenna port and 64 accumulated SFs, the RC ML algorithm reaches a 10% FO hit rate at an SNR of −16.5 dB for the EPA model and −17.3 dB for the ETU one, instead of −21.6 dB for the static AWGN case. Even though the difference is approximately 3 dB smaller than for static channels, the ML RC algorithm still clearly outperforms the PD approach by at least 8.5 dB for FO estimation for the two fading channels. As for the TO estimation, the conclusions are similar for the EPA channel. But for the ETU channel, the two approaches have a high residual miss rate. With an RMS delay spread of approximately 2 samples, ETU channels have strong taps with large delays that can have higher power than the first tap for some realizations. In this case, first path detection is needed to discriminate

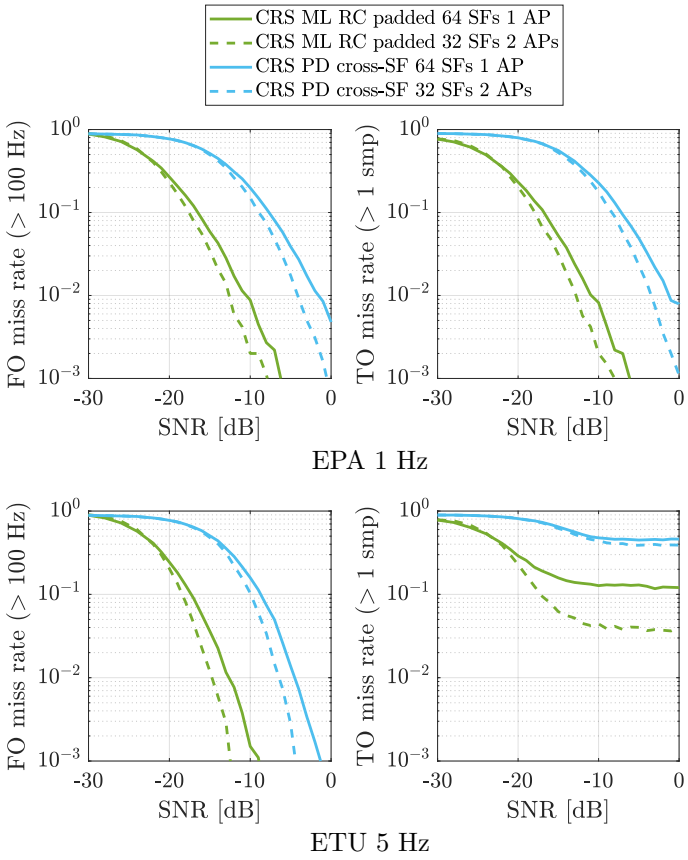

Figure 3.18: FO (left) and TO (right) estimation miss rates without hit detection for the ML RC padded and PD estimators with CRS in EPA 5 Hz (top) and ETU 1 Hz (bottom) fading channels. Simulated with 10000 iterations per SNR point.

the different channel taps and improve the estimation of the time offset.

If the cell uses multiple antenna ports, the TO estimation can be improved by taking advantage of diversity. The estimation performance for two CRS APs is also depicted in Fig. 3.18. An accumulation time of 32 SFs was chosen to evaluate the results with the same received energy as with a single antenna port and 64 SFs. For the FO estimation, the two methods can exploit the diversity gain and improve performance by approximately 1 dB. But for the TO estimation, the residual miss rate is reduced to 4% with the ML RC algorithm while it remains in the order of 40% for the PD approach.

The results for the combined FO and TO miss rate and including the hit detection for the ML RC algorithm are shown in Fig. 3.19. Hit detection comes at the cost of 1.8 dB at a 10% miss rate, but allows the correction of wrongly detected frequency offsets to be skipped. Overall, the target −18.5 dB SNR mentioned in the MCL evaluations of Chap. 2 can already be achieved with 128 SFs for the EPA channel,

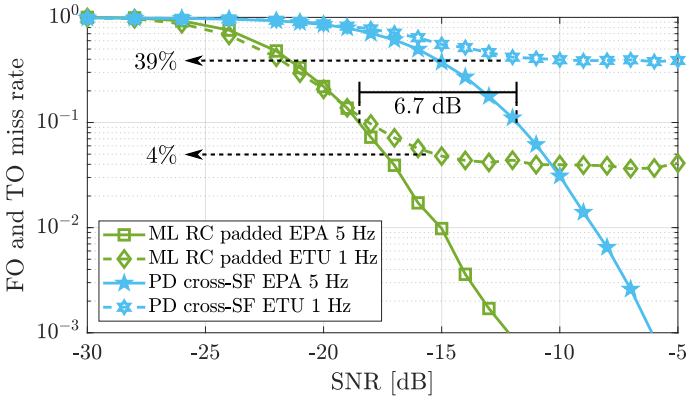

Figure 3.19: FO and TO estimation miss rate for the ML RC padded and PD estimators with and without hit detection, respectively. The two methods use 128 SFs and 2 APs. Simulated with EPA 5 Hz (solid lines) and ETU 1 Hz (dashed lines) channels and 10000 iterations per SNR point.

while higher accumulation numbers are possible with the packed version of the algorithm at the cost of a reduced estimation range. On the other hand, the ETU channel does not only cause a residual miss rate of 4%, but it also produces a similar rate of false time offset hits that could lead to wrong corrections.

In future work, some of the intermediate results of the ML RC algorithms could be reused to improve the TO estimation performance with limited impact on computational complexity. The estimation metric of the discrete time offset $\hat{u}$ described in (3.23) of Sec. 3.5.3 could be used for first path detection in a post-processing step. The same metric could also be used to generate additional conditions for the hit detection, instead of relying only on the FO metric, to filter out the miss detected time offsets. If needed, the range of the TO estimation could also be extended by skipping the combination of CRS REs belonging to the same OFDM symbol described in Sec. 3.5.2, doubling the memory requirement and the complexity of the algorithm. Finally, the performance of the FO estimation for the high mobility channels with up to 200 Hz Doppler frequency specified for eMTC conformance testing in LTE Release 15 [16] should be assessed.

3.6.3 Summary

The different options for fine synchronization evaluated in this work or found in the literature and listed in Sec. 3.4.1 are:

- *CP AC*: cyclic-prefix autocorrelation in the time domain [41,42].
- *PSS-SSS ML*: maximum likelihood estimation with PSS and SSS cross-correlations [13].
- *CRS PD*: phase difference between CRS [39, 46].
- *PBCH*: phase difference between PBCH repetitions [33].
- *CRS ML RC*: reduced complexity maximum likelihood estimation with CRS presented in this work.

The properties of these estimators are summarized in Fig. 3.20 for a qualitative comparison. The methods are based on their FO estimation features, following the conclusions of [33] and the evaluations

of this chapter. The axes are labelled to represent better scores on the outer ticks, e.g. computational complexity has the lowest values at the boundary. The focus is set on the FO estimation because it usually limits the synchronization performance in eMTC and because not all approaches include TO estimation.

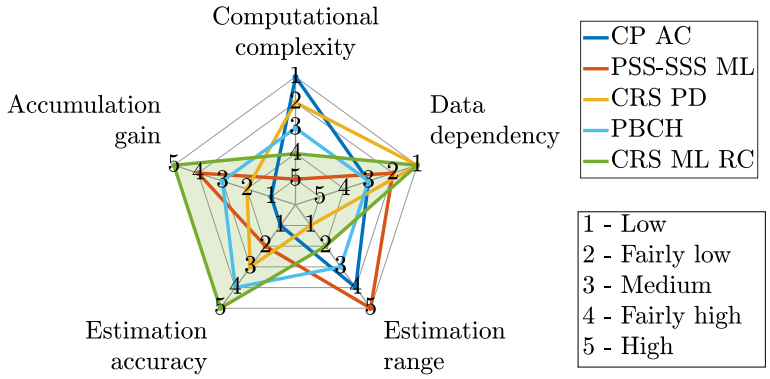

Figure 3.20: Qualitative comparison of fine FO estimation methods.

The CP AC, CRS PD, and PBCH algorithms have the lowest computational complexities. However, for all three, increased accumulation time yields diminishing returns at low SNR due to the noise enhancement effect discussed for the CRS PD method. Besides its complexity, the main drawback of the PSS-SSS-ML fine synchronization is its low FO estimation accuracy. Even at moderate SNR levels accumulation over multiple signal periods is required to cope with the reduced length of the correlation sequence and reach RMSE levels below 100 Hz.

The presented CRS ML RC estimator is the only solution that yields sufficient accumulation gain and FO estimation accuracy to fulfil the requirements of extended coverage operation in cellular IoT networks. Furthermore, thanks to the availability of CRS, this method applies to all cell configurations, does not depend on cell traffic, and can be used throughout the whole cell bandwidth, the criteria used for the data dependency ranking of Fig. 3.20. The two main drawbacks of this approach were addressed by applying techniques to reduce

computational complexity and to achieve a sufficient FO estimation range. Therefore, the CRS ML RC algorithm was selected for eMTC fine synchronization in the developed modem SoC. The relevant implementation details and the performance measured on the fabricated chip are presented in subsequent chapters.

Chapter 4

OTDOA Positioning[1]

Observed time difference of arrival (OTDOA) positioning capabilities in LTE were introduced with 3GPP Release 9 in 2009 as a tool to comply with location accuracy requirements of emergency call systems like enhanced 911 (E-911) [28]. Targeting an accuracy of approximately fifty to a hundred meters, OTDOA has not yet established itself in other mainstream applications where GNSS positioning is dominant. The common OTDOA procedure requires timing range measurements to be sent to a location server where the actual position is calculated and sent back to the UE if it needs the information, inducing additional data transfers and privacy concerns in most human-centric applications where the location information is mostly used locally. However, the rise of the IoT has paved the way for use cases that do not require a high precision position estimation and where the information is not needed by the device itself but uploaded to a server.

In this context, the 3GPP made a point of maintaining OTDOA support in its cIoT standards. As a standalone solution, cellular positioning can provide a ubiquitous and low-cost alternative in case of widespread adoption, benefiting from the global availability of cellular networks and the relatively small overhead required to upgrade UEs with OTDOA capabilities, compared to adding the support of other positioning technologies. When used as a complement to GNSS,

[1]Part of this chapter was adapted from [50] published in 2020 IEEE/ION PLANS.

OTDOA can help expand location coverage to indoor and urban scenarios, and improve the robustness of the system against attacks by adding redundancy to the positioning information.

The focus of this chapter is the theoretical and algorithmic background to support OTDOA positioning, while SoC implementation and measurements are discussed in subsequent chapters. First, the general OTDOA positioning procedure and working principle are introduced in the scope of NB-IoT and eMTC. Then, the so-called positioning reference signals (PRS), designed to improve the hearability of neighbour cells, are presented before discussing the TOA measurement itself, the core part of the positioning procedure. Finally, the effect of frequency offset on the coherent accumulation of PRS subframes is investigated, showing the importance of fine synchronization.

4.1 OTDOA in Cellular IoT

OTDOA is a DL positioning method where the location of the UE is estimated using pseudo-range multilateration from different base stations. The pseudo-ranges are obtained with measurements performed by the UE on the signals received from different BSs. Assuming synchronized transmitters, the difference in the TOA of the signals is caused by their distinct propagation delays, which are proportional to the distances between the UE and the base stations. Only differential measurements of the propagation delays are possible for the UE because it does not have any information on the absolute transmission time of the signals. The differential time delay measurements are obtained by subtracting the estimated TOA of the signal of a reference cell from the TOAs of the signals transmitted by neighbour cells. Converting the time measurements into distance differences, the possible positions of the UE referenced to a pair of neighbour and reference cells are given by a hyperbola as illustrated in Fig. 4.1. A planar position fix can be obtained with at least two range measurements and is graphically represented by the crossing of two hyperbolas. If more neighbour cells are available, the redundant information can be used to improve the location accuracy.

OTDOA positioning in cellular IoT systems is similar to its LTE counterpart. For eMTC, OTDOA support was inherited from LTE

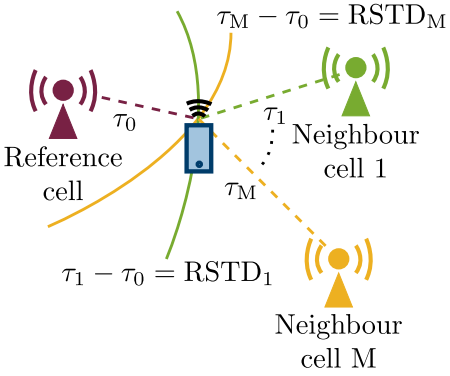

Figure 4.1: Multilateration and OTDOA working principle based on RSTD measurement.

but enhanced in a second standardization stage with improvements in Release 14 [51], like adding performance requirements and increasing the repetitions of the available reference signals for eMTC devices. As to NB-IoT, OTDOA positioning was not supported in its first version, but the standard was upgraded to support OTDOA in Release 14 [52].

NB-IoT and eMTC reuse the network architecture of LTE where location-based services (LBS) are managed by a location server [7]. In this setting, the interactions and communication between a UE and its location server are governed by the LTE Positioning Protocol (LPP) [53]. The different steps of the common OTDOA positioning procedure are depicted in Fig. 4.2. After establishing a connection and registering to the network, the UE informs its location server about its positioning capabilities. Then, the UE is provided with the assistance data needed for the time difference of arrival (TDOA) measurements, like the relevant configuration of the reference and neighbour cells that should be used, and expected values of the time differences to reduce the search space that must be covered. With this information the UE can perform the actual TDOA measurements, also called reference signal time difference (RSTD) measurements, and include the results in a location information transfer sent to the server that estimates the position of the device.

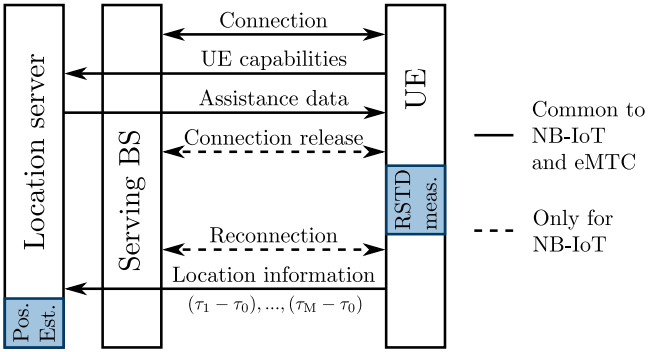

Figure 4.2: OTDOA procedure [50].

The two main challenges faced to support OTDOA in cIoT networks are the targeted coverage extension and the limited capabilities of devices [54]. Especially, NB-IoT was designed to restrict the number of connected mode procedures to reduce the number of parallel tasks performed by UEs and minimize device complexity. Therefore, RSTD measurements are performed in the idle state. The connection of the UE has to be released and reestablished before and after the measurements, as depicted by the arrows with dashed lines in Fig. 4.2. For eMTC and LTE, RSTD measurements belong to connected mode procedures. The reduced RX bandwidth of NB-IoT and eMTC UEs impacts the achievable accuracy of RSTD measurements and has to be considered, along with the reduced signal-to-interference-plus-noise ratio (SINR) in extended coverage scenarios, in the definition of minimum measurement accuracy and reporting delay requirements. The accuracy requirements of intra-frequency RSTD measurements and the assumed operating conditions are summarized in Tbl. 4.1 as specified by the 3GPP [29]. The UE must fulfil the defined requirements in 90% of measurement iterations. Two different LTE scenarios are considered in the table, one assumes a 1.4 MHz cell and the other a 5 MHz one, both with reference signals occupying the full resource grid in the frequency domain. For NB-IoT and eMTC two scenarios are included, normal coverage (NC) with similar SINR conditions for the reference (ref.) and neighbour (neigh.) cells as

	LTE		eMTC		NB-IoT	
	5 MHz	1.4 MHz	NC	EC	NC	EC
Signal BW [MHz]	4.5	1.08	1.08	1.08	0.18	0.18
SF number	2	6	12	30	320	320
Ref. SINR [dB]	-6	-6	-6	-15	-6	-15
Neigh. SINR [dB]	-13	-13	-13	-15	-13	-15
Accuracy [ns]	195	488	488	488	651	1040
Accuracy [m]	59	146	146	146	195	312

Table 4.1: Summary of operating conditions and RSTD measurement requirements for OTDOA positioning support [29].

LTE, and extended coverage (EC) with more challenging conditions. The relevant parameters for the reference signals are their bandwidth (BW) and the number of SFs that are available for the measurements. The accuracy requirements are specified in LTE 30.72 MHz samples but were translated into nanoseconds and meters in the table for easier interpretation.

The table shows that the accuracy requirement was relaxed and the number of measurement subframes increased for decreasing bandwidth and SINR. For narrowband signals, the bottleneck of the positioning performance is the RSTD measurement, even though the synchronization between BSs as well as the size and geometry of the constellation of visible cells also have an impact. The RSTD measurement is similar to the time offset estimation part of the problem discussed in Chap. 3, but applied to multiple cells. In this case, the lower bound on the mean squared error (MSE) of the TO estimate was shown to be inversely proportional to the signal bandwidth cubed, i.e. the number of REs carrying reference signals K in (3.10), but only inversely proportional to the number of SFs n_{SF}. In other words, narrowband signals offer less TOA estimation accuracy and require longer accumulation times to compensate for the loss due to their reduced bandwidth. In principle, any reference signal could be used for the RSTD measurement. However, to improve the quality of the measurements, dedicated reference signals were introduced specially for OTDOA. Their structure and properties are discussed next.

4.2 Positioning Reference Signals

For eMTC, positioning reference signals (PRS) were inherited from LTE, while for NB-IoT a new set of reference signals had to be introduced, namely narrowband PRS (NPRS). In both cases, a pseudo-random sequence of QPSK symbols is mapped every six resource elements on symbols that do not carry other mandatory channels or signals, such as synchronization signals, broadcast channels, or legacy LTE control channels. The sequence of QPSK symbols depends on the cell-specific PRS ID and the SF number of the transmitted signals. When represented on the OFDM resource grid, the mapping follows a diagonal pattern as shown in Fig. 4.3. PRS and NPRS have separate antenna ports and are only mapped on specific subframes. For LTE and eMTC, the set of SFs containing PRS is defined by the periodicity of the mapping, the offset of the starting position within that period, and the number of consecutive subframes that contain PRS, building a positioning occasion. One of the improvements introduced for eMTC allows the definition of more flexible mappings, e.g. by increasing the maximum length of a positioning occasion to 160 instead of 6 SFs. For NB-IoT, two mapping definitions are possible, and they can be used standalone or combined. One option also uses periodicity, starting position, and length of a positioning occasion, while for the other one the valid PRS subframes are defined with a bitmask of length 10 or 40, where each bit indicates whether the corresponding SF contains NPRS. In NB-IoT, a positioning occasion can have up to 1280 consecutive SFs, allowing longer accumulation times.

Positioning occasions consist of low interference subframes, without data transmission on any cell to make sure that positioning signals are not covered by other channels. To further reduce inter-cell interference on positioning occasions, selected positioning SFs can be transmitted with zero power, i.e. muted. The muting pattern is chosen by the network operator to fit its deployment. Overall, PRS and NPRS have three layers of isolation. In the frequency domain, the mapping to resource elements allows a reuse factor of six with low interference as long as the orthogonality of subcarriers is maintained, i.e. the FO of the UE and the TO between the different cells are small enough, in the order of a few hundred hertz and tens of microseconds,

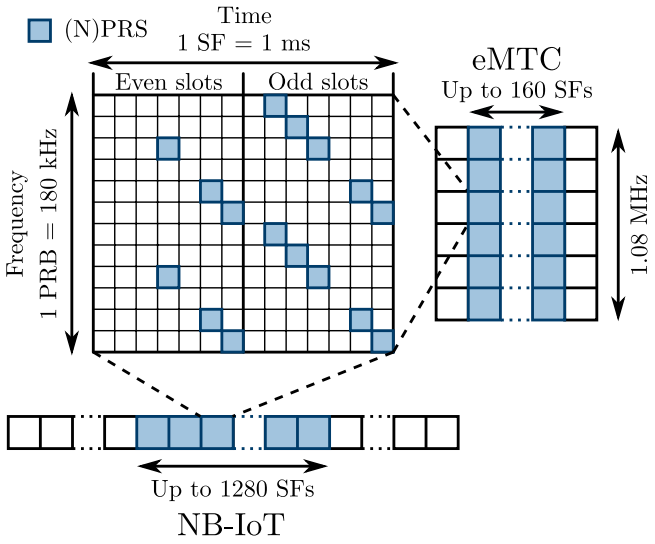

Figure 4.3: Example of OTDOA PRS and NPRS mapping [50].

respectively. In the time domain, isolation can be introduced with muting. Finally, PRS-ID-dependent sequence generation ensures that the signals can still be separated in the code domain if they overlap in time and frequency.

4.2.1 CRLB Comparison

The CRLB on the TOA estimation with PRS and NPRS allows the comparison of the accumulation time needed to achieve the RSTD measurement accuracy requirements stated in Tbl. 4.1 for eMTC and NB-IoT. Recalling the signal model of (3.2) for a static AWGN channel, without frequency offset, and assuming only the i^{th} cell with a delay τ_i is active, the received signal for the l^{th} symbol is :

$$r_l(nT_s) = s_l^{(i)}(nT_s - \tau_i) + w_l(n) \ . \tag{4.1}$$

For eMTC, the baseband signal modulation is described in (3.1). NB-IoT has a similar signal with a half-subcarrier shift and only the $N_{\mathrm{sc}} = 12$ center subcarriers occupied [18]:

$$s_l^{(i)}(nT_s - \tau_i) = \frac{1}{\sqrt{N_{\mathrm{FFT}}}} \sum_{k=-N_{\mathrm{sc}}/2}^{N_{\mathrm{sc}}/2-1} X_{k,l}^{(i)} e^{2\pi j(k+\frac{1}{2})(nT_s-\tau_i)\Delta f} \ . \quad (4.2)$$

With these signal definitions, assuming a noise term with variance σ^2 and unity power modulation symbols $X_{k,l}^{(i)}$, the SNR per RE is equal to $\frac{1}{\sigma^2}$, which is the standard SNR definition assumed throughout this work, and corresponds to the one used for the 3GPP's normative requirements. Furthermore, the CRLB of the TOA τ_i for eMTC is the same as for LTE given in [28]. For NB-IoT only a small adaption is needed to account for the half subcarrier shift [50]. For a single subframe and without taking into account the CP the CRLB is:

$$\mathrm{CRLB}(\hat{\tau}_i) = \frac{\sigma^2}{8\pi^2\Delta_f^2 \sum_{l=0}^{N_{\mathrm{symb}}-1} \sum_{k=-N_{\mathrm{sc}}/2}^{N_{\mathrm{sc}}/2-1} (k+\delta_k)^2 |X_{k,l}^{(i)}|^2} \ , \quad (4.3)$$

where $\delta_k = 0$ for eMTC and $\delta_k = \frac{1}{2}$ for NB-IoT. When accumulating multiple SFs the achievable estimation variance is divided by the number of combined subframes n_{SF} [50]:

$$\mathrm{CRLB}(\hat{\tau}_i, n_{\mathrm{SF}}) = \frac{\mathrm{CRLB}(\hat{\tau}_i)}{n_{\mathrm{SF}}} \ . \quad (4.4)$$

Thus, the ratio of the CRLB for NB-IoT and eMTC can be used to compare the number of subframes required to achieve similar performance in both. Assuming the same SNR per RE and the mapping depicted in Fig. 4.3 for both, 221 times more SFs are needed in NB-IoT:

$$\frac{n_{\mathrm{SF,NB\text{-}IoT}}}{n_{\mathrm{SF,eMTC}}} = \frac{\mathrm{CRLB}_{\mathrm{NB\text{-}IoT}}(\hat{\tau}_i)}{\mathrm{CRLB}_{\mathrm{eMTC}}(\hat{\tau}_i)} = 221 \ .$$

When assuming the same received power instead of the same SNR per RE the difference reduces to approximately 37 times. These results have also been shown in [50] where a compensation term for

the different number of occupied subcarriers is taken into account. The difference in TOA estimation performance is due to the smaller signal bandwidth supported in NB-IoT and is compensated with up to 26 times longer accumulation times compared to eMTC. However, the number of available SFs is not sufficient to fully recover the estimation performance, and the target accuracy requirements had to be relaxed to avoid excessively long accumulation times.

4.3 TOA Estimation

The TOA of the signal transmitted by the i^{th} cell can be estimated using the correlation of the received signal $r_l(n)$ with a replica of the original signal $s_l(n)$. The correlation function over one SF for a given time delay τ can be calculated as the sum of symbol correlations in the time domain:

$$R_m^{(i)}(\tilde{\tau}) = \sum_{l=0}^{N_{\text{symb}}-1} \sum_{n=-N_{\text{CP},l}}^{N_{\text{FFT}}-1} r_l(n+\tilde{\tau})(s_l^{(i)}(n))^* \ .$$

The ML TOA estimation in AWGN is then given by the delay that maximizes the correlation power for a given search window defined by $-W \leq \tilde{\tau} < W$ [28]:

$$\hat{\tau}_i = \arg\max_{\tilde{\tau}} \left| R_m^{(i)}(\tilde{\tau}) \right|^2 \ .$$

Excluding noise, interference, and cross-correlation from adjacent SFs, the correlation output with zero time delay is the power of the autocorrelation function of the PRS or NPRS subframe, both plotted in Fig. 4.4 for 1.92 Msps sampling rate. Due to their smaller bandwidth, NPRS yield a much wider and flatter peak, making the estimation of its location less accurate in noisy conditions, as shown by the CRLB comparison.

The estimator can be extended to use n_{SF} subframes by accumulating their correlations. The cross-subframe combination can be coherent, adding the cross-correlations before calculating the power of the accumulated values, or non-coherent, adding the power of the

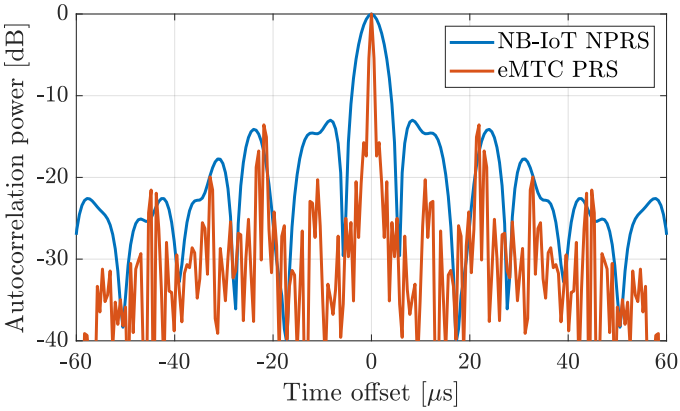

Figure 4.4: PRS and NPRS autocorrelation power at 1.92 Msps. The values are normalized to the peak of each curve.

correlation of each SF. The correlation function can be modified to coherently accumulate blocks of n_c SFs, while combination across blocks is non-coherent, to obtain the following estimation metric also presented in [50]:

$$R(\tilde{\tau}) = \sum_{p=0}^{n_\mathrm{b}-1} \left| \sum_{m=0}^{n_\mathrm{c}-1} R_{pn_\mathrm{c}+m}(\tilde{\tau}) \right|^2 \quad , n_\mathrm{b} n_\mathrm{c} = n_\mathrm{sf} \ . \tag{4.5}$$

This formula applies in general for OTDOA estimation with multiple received subframes. The remainder of this section will focus on the special case of NB-IoT.

4.3.1 Quadratic Interpolation

The lower sampling rate of NB-IoT and eMTC devices reduces the complexity of the cross-correlation used to estimate the TOA but comes at the cost of a reduced granularity of the estimation metric. To restore resolution, an additional interpolation stage is required after performing the maximum detection on the correlations of (4.5).

A simple quadratic fitting scheme was chosen instead of the sinc-based interpolation described in [55] to keep the complexity as low as possible with a negligible impact on the interpolation accuracy. With the mentioned sampling rate, the NB-IoT signal is still oversampled by roughly a factor of ten. Thus, the wide peak of $R(\tau)$, shown in the NPRS autocorrelation of Fig. 4.4, can be well approximated by a quadratic function. The interpolated maximum can be obtained by fitting a parabola on the maximum bin of $R(\tau)$ and its two neighbours. The estimated TOA is then given by [50]:

$$\Delta\hat{\tau} = \frac{R(\hat{\tau}+1) - R(\hat{\tau}-1)}{2(2R(\hat{\tau}) - R(\hat{\tau}+1) - R(\hat{\tau}-1))} \ . \tag{4.6}$$

4.3.2 Interference from Neighbour Cells

Under static AWGN channel conditions and without synchronization mismatches, the coherent combination yields better performance than the non-coherent because it allows ML estimation. Furthermore, coherent accumulation is also more resilient to the interference of PRS from neighbouring cells. When estimating the TOA of a given cell, e.g. the reference cell, the interference from other cells results in an unwanted cross-correlation term in the estimation metric of (4.5). If the interfering cell is much stronger than the wanted cell, the peak of the interfering term may exceed the peak of the autocorrelation function of the wanted signal, leading to the miss-detection of the actual maximum. For NB-IoT, the cross-correlation of NPRS from a reference cell with NPRS ID 0 and a neighbour cell with NPRS ID 6 is depicted in Fig. 4.5. The NPRS of the two cells are mapped to the same subcarriers, leading to high cross-correlation values at low time offsets, where the signals are only separated in code domain by the pseudo-random sequence of QPSK symbols that build the reference signals. The correlation power for $n_{\mathrm{sf}} = 4$ consecutive SFs with the coherent accumulation of blocks of one, two or four SFs are shown. The values are normalized to the peak power of the autocorrelation of the serving cell with the same accumulation strategy, i.e. 0 dB would mean that the interfering cross-correlation has the same power as the wanted peak. The dotted lines highlight the maximum of each curve, which represents the margin of the wanted peak on the interference

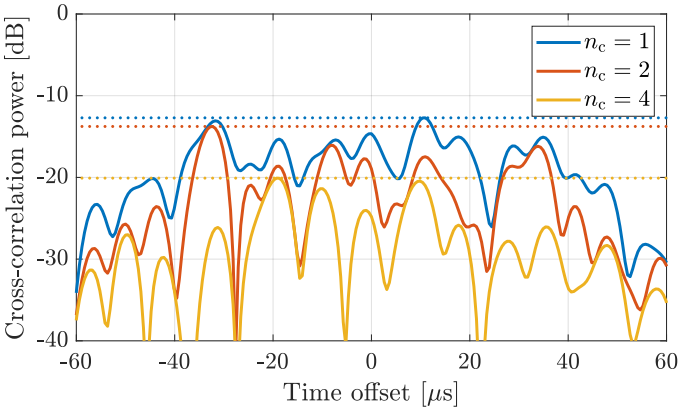

Figure 4.5: NPRS cross-correlation power of neighbour and reference cell signals accumulated over $n_{sf} = 4$ SFs with different numbers of coherently accumulated SFs. The values are normalized to the peak of the reference cell's autocorrelation for the same configuration.

term. If the maximum of a curve in Fig. 4.5 is at −12 dB, the TOA of the wanted cell can be correctly detected even if its power is up to 12 dB lower than the neighbour cell. In the illustrated example, increasing the number of coherently accumulated SFs helps reduce the interfering peak's amplitude compared to the wanted maximum.

The amplitude of the interfering cross-correlation and the benefits of coherent combination depend on the sequences used to generate the NPRS and on their mapping to resource elements. To verify that coherent accumulation generally increases resilience to interference and to quantify this improvement, the maximum cross-correlation value was calculated for all possible neighbour NPRS IDs. The maximum cross-correlation power relative to the corresponding autocorrelation peak averaged over all IDs is shown in Tbl. 4.2. When accumulating 320 NPRS SFs, coherent accumulation of all SFs allows on average 2.6 dB higher interferers than coherently accumulating groups of two SFs, and another 2.1 dB higher than the non-coherent accumulation of SFs, confirming the benefits of the coherent combination shown

Number of coherently accumulated subframes n_c	Average maximum cross-correlation
1 SF	−13 dB
2 SFs	−15.6 dB
4 SFs	−17.7 dB

Table 4.2: Average maximum NPRS cross-correlation power over 160 accumulated SFs for a reference cell with NPRS ID 0 and all possible neighbour cell IDs. The power value is given in decibel relative to the autorcorrelaion peak of the reference cell.

in the example of Fig. 4.5. For similar reasons, coherent combination across subframes can help mitigate interference also for other reference signals that rely on separation in code domain, like NRS, CRS or DMRS.

However, this advantage relies on the phase coherency between subsequent subframes and is vulnerable to residual frequency offsets and changes in the radio channel. Thus, the number SFs n_c that can be coherently accumulated without losing performance is limited by frequency mismatches and channel Doppler spread. Successive interference cancellation, as presented in [55], can also improve performance in interference-limited scenarios. However, the achieved improvements come at a high computational cost and the extension of this approach for large numbers of accumulated subframes for coverage extension would significantly increase its complexity. Therefore, it was not considered a viable solution for the developed cellular IoT SoC.

4.4 NB-IoT Performance Evaluation

The performance of the described TOA estimation algorithm was evaluated with Monte-Carlo simulations, assuming an AWGN channel as defined in (4.1) and with 2000 channel realizations for each SNR point. To reflect the conditions of a real-world scenario, the virtual positions of the UE and the simulated cells were chosen based on a real location at the institute and the position of four neighbouring base stations, corresponding to an irregular urban deployment with

	Distance to UE	Distance difference	RX power difference
Reference cell	340 m	0 m	0 dB
Neighbour cell 1	691 m	351 m	−11.6 dB
Neighbour cell 2	615 m	273 m	−9.7 dB
Neighbour cell 3	882 m	546 m	−15.6 dB

Table 4.3: The geometrical arrangement of cells for OTDOA simulations and resulting signal parameters [50].

an inter-cell distance of up to 1 km. Tbl. 4.3 summarizes the distances between the UE and the base stations, as well as the resulting power ratios and relative time delays of the signals received from different cells. The power differences were calculated based on the distance differences using the same path loss model as in [55]:

$$L[\text{dB}] = 120.9 + 37.6 \log_{10}(d[\text{km}]) \ .$$

The closest cell was chosen as the reference cell, while the three others were defined as neighbour cells. The RSTD for each neighbour cell was calculated as the difference between the TOA of its signal and the TOA of the reference cell's signal, where the TOA was measured as described in previous sections, with $n_c = 4$ coherently accumulated subframes. The NB-IoT cells were set to operate in standalone mode with 320 NPRS SFs available for the RSTD measurement, without muting. The cell IDs were set such that the NPRS of all cells were mapped on different subcarriers. Fig. 4.6 shows the miss rate of the simulated measurements for each neighbour cell. The measurement was considered a miss when the absolute error $|\mathcal{E}_{\text{RSTD}}|$ compared to the expected RSTD value was larger than 651 ns, i.e. the accuracy requirement for normal coverage defined in Tbl. 4.1. To ensure a sufficient margin on the specification, a miss rate of 5%, highlighted by the black dotted line in Fig. 4.6, was considered instead of the normative 10%.

When expressed relative to the SNR of the corresponding neighbour cell instead of the reference cell, the three curves show almost the same performance, achieving a 5% miss rate at SNRs between

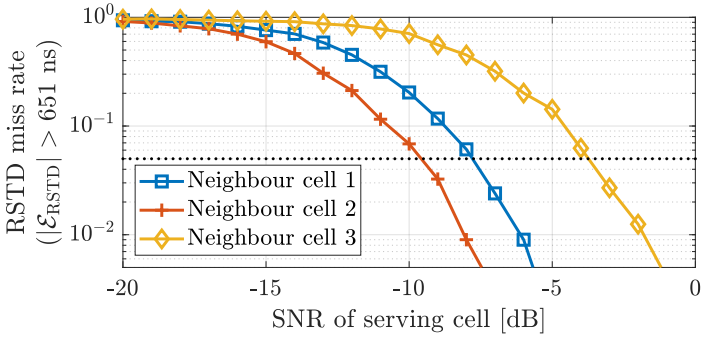

Figure 4.6: RSTD measurement miss rate in a static AWGN channel for 320 accumulated NPRS subframes. Simulated with 2000 iterations per SNR point.

-19.3 and -19.4 dB. Thus, in AWGN conditions, the performance of the RSTD measurement has a margin of at least -4.3 dB on the extended coverage SNR requirement for the three neighbour cells. However, this result can be influenced by other non-idealities that affect the behaviour of a real UE. In particular, the residual FO can heavily impact the TOA estimation performance when NPRS SFs are accumulated coherently. This effect is investigated next.

4.4.1 Effect of Frequency Offset

To assess the impact of a residual frequency offset (FO) of f_{o} Hz on the RSTD measurement, the model of (4.1) was extended [50]:

$$y_l(n) = s_l(n - \tau_i)e^{2\pi j n T_s f_{\mathrm{o}}} + w_l(n) \ . \tag{4.7}$$

The simulations of Fig. 4.6 were repeated with this additional effect for different numbers of coherently accumulated SFs $n_{\mathrm{c}} \in \{1, 2, 4\}$ and still a total of $n_{\mathrm{sf}} = 320$ NPRS SFs. To emulate the behaviour of real systems, where the local oscillator's frequency varies in time, the FO was assumed to be equally distributed in $]-F_{o,\max}, F_{o,\max}[$ for different values of $F_{o,\max}$. The results are plotted in Fig. 4.7, where the maximum FO value for a given simulation point is represented on

the horizontal axis. The vertical axis indicates the 5 % RSTD miss rate SNR point of the strongest neighbour cell, i.e. neighbour cell 2 in Tbl. 4.3. The SNR is relative to the considered neighbour cell.

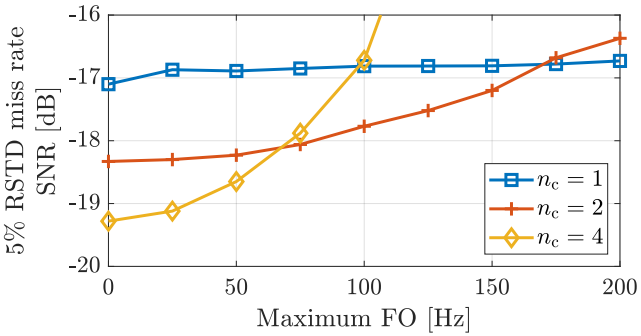

Figure 4.7: Effect of frequency offset on RSTD measurement performance [50]. Simulated in a static AWGN channel for 320 accumulated NPRS subframes and with 2000 iterations per SNR point.

As expected, at zero FO the correlations with more coherently accumulated SFs perform better, indeed considering the origin of the x-axis $n_c = 4$ performs best. However, its performance degrades quickly with increasing FO. FOs in the order of 50 Hz to 100 Hz are common in real-world scenarios due to the temperature-sensitive low-cost oscillators used in cIoT modems. Thus, the benefits of coherent accumulation can only be realized with accurate frequency synchronization for which solutions have already been presented.

Time-varying fading channels also limit the coherency of reference signals over time. The maximum Doppler frequency of a channel influences the number of subframes that should be combined coherently. Furthermore, multipath channels introduce uncertainty in the measured TOA and considerably impact the accuracy of RSTD measurements. First-path detection helps reduce the degradation due to fading channels at the cost of additional post-processing on the time-domain correlations of (4.5). The effects of multipath on OTDOA positioning and the possible solutions to tackle them were not studied

in the scope of this thesis. Nevertheless, they can be implemented incrementally on the presented measurement strategy in future work.

Chapter 5

System-on-Chip Architecture and Operation

A high integration level is crucial to meet the stringent requirements of mIoT applications in terms of cost, form factor, and power consumption. In this setting, RF SoCs, including analog radio frequency transceivers and processor systems, are the best-suited platforms to meet the needs of a variety of use cases and increase the number of deployed devices. The benefits of integration outweigh the inconvenience of coupling the technology used for the RF and digital sub-systems, which may slow the migration of the latter to newer processes.

Processor systems typically included in such SoCs do not provide sufficient resources for a purely software-defined radio (SDR) implementation of the digital PHY, even for the reduced complexity of cIoT standards. Moreover, coverage extension increases the complexity of algorithms needed to set up and maintain a connection, and to support additional services like OTDOA positioning. Therefore, hardware acceleration is required, at least for a subset of the signal processing tasks. Tight integration of digital baseband blocks in the processor system enables fine partitioning and smooth transitions between software and hardware, allowing trade-offs between the advantages

of hardware acceleration, regarding power consumption and performance, against the additional silicon cost of dedicated processing blocks. Furthermore, flexible and modular systems are highly beneficial to facilitating the incremental development of multimode platforms supporting various communication and positioning technologies with shared resources.

In this chapter, the architecture of state-of-the-art modem SoCs supporting NB-IoT and eMTC is presented, highlighting the need for hardware acceleration in digital signal processing, and presenting trade-offs in HW-SW co-design. The performance requirements that drive the digital system design are investigated first. Then, a system overview of the modem SoCs where the developed signal processing blocks were integrated is given. Following, the digital hardware involved in the implementation of tracking for eMTC and in OTDOA measurements for NB-IoT positioning is described along with the interactions between HW and SW.

5.1 System Requirements

The choice in partitioning between HW and SW processing depends mainly on the complexity of the tasks and the time available to perform them. Even if a dedicated hardware implementation provides better power efficiency for specific functionality, the benefits are attenuated when the power consumption of the whole system is considered. Many mIoT use cases are characterized by sporadic activity dominated by UL data traffic. During active bursts, energy consumption is mainly driven by the PA, while sleep power is the most relevant during idle periods. Instead, using SW where possible and building accelerators that can be reused for different tasks is preferred for low-cost devices to minimize area.

The digital signal processing steps directly involved in the reception and transmission of physical channels are summarized in Fig. 5.1. Operations like OFDM demodulation in the RX path and signal generation in the TX chain must operate at the sampling rate, here 1.92 Msps. Generally, they need dedicated hardware because they require online processing to avoid buffering large amounts of samples and because they have strict deadlines imposed by the standard. At

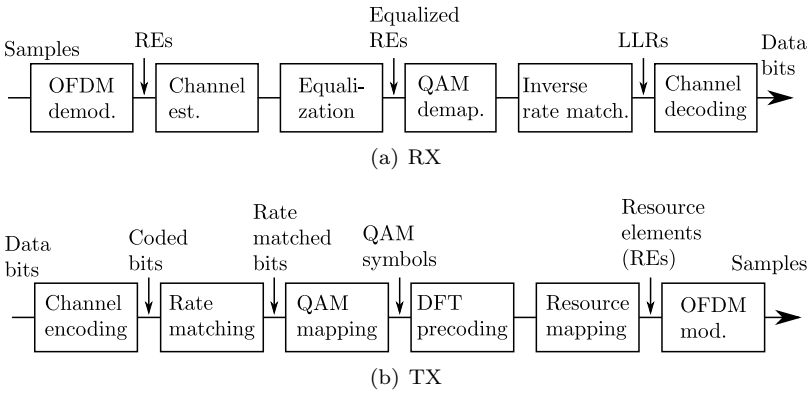

Figure 5.1: Digital signal processing steps of RX (a) and TX (b) chains.

the other end of the chain, bit operations like channel encoding for signal generation can be implemented very efficiently in HW, while channel decoders in the RX path are too complex to be implemented in SW and meet the latency and throughput requirements for both NB-IoT and eMTC. Whether the intermediate steps, like channel estimation and equalization for the DL and DFT precoding for the UL, can be implemented in SW depends on the number of REs that must be processed per symbol and on the available time to complete the operations. The timing relations between control and data channels defined by the standard impose latency constraints on these tasks.

Fig. 5.2 shows the minimum delays between control and data channels in the transmission and reception of data in NB-IoT and eMTC. Besides having a six times smaller bandwidth, NB-IoT also has latencies between three and four times longer. The combination of these properties substantially relaxes the computational requirements of NB-IoT, allowing more operations to be performed in software. For example, the channel estimation on REs that contain reference signals and the equalization of REs that carry control information must be performed approximately 24 times faster in eMTC compared to NB-IoT. This number holds if the time allocated to each stage of the RX path is proportionally the same for the two standards

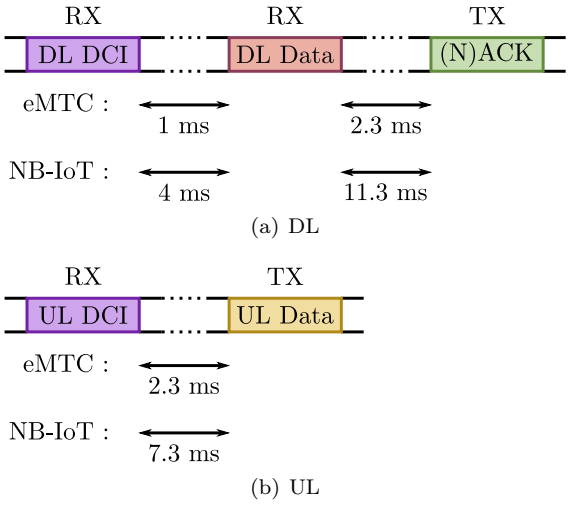

Figure 5.2: Timing diagram of the dynamic scheduling of DL (a) and UL (b) data channels.

and if in both cases the performance is constrained by latency and not throughput. A factor of four comes from the difference in the 1 ms delay between MPDCCH and PDSCH and the 4 ms between NPDCCH and NPDSCH, and a factor of six is due to the bandwidth difference resulting in a larger number of REs processed per subframe. Furthermore, for eMTC, the MPDCCH carries the narrowband index of the NPDSCH. Thus, the control channel must be fully decoded early enough to allow a frequency hop in time to decode the scheduled PDSCH. The increased complexity combined with the strict deadline requires all the RX steps involved in MPDCCH decoding to be performed in HW for eMTC while some of them could be shifted to SW for NB-IoT.

The small latency between control and data channels in uplink scheduling is also challenging because it involves fetching data from higher protocol layers besides the full completion of all the steps depicted in Fig. 5.1 of both RX and TX paths. Again, the lower

latency and higher complexity of eMTC make HW accelerators needed for the whole RX and TX chains.

Coverage extension support also influences the computational requirements of the system. On the one hand, operation at low SNRs impacts the complexity of the algorithms needed for synchronization and data transmission. On the other hand, real-time processing is needed to avoid the usage of large memory buffers for offline processing. If no explicit deadline is imposed by the standard, the complexity of the algorithm and the signal periodicity determine the computational requirements. For example, there is no stringent requirement on the time needed for cell search. However, coverage extension needs ML detection with a large number of signal periods. To allow online accumulation, the needed cross-correlations must be computed within a signal period and need a high-throughput hardware implementation as described in [13].

5.2 Overall Architecture

Fig. 5.3 shows an overview of the architecture of the multimode SoCs presented in [56] and [13], for which contributions were done in the scope of this work. The considered systems include RF and central processing unit (CPU) subsystems, as well as digital accelerators for compute-intensive tasks.

The RF transceiver subsystem, in the top-left part of the block diagram, was designed by a partner company. Its RX chain includes LNA, down-mixer, filtering, and ADC stages. It supports all GSM and LTE frequency bands up to 2.2 GHz, including those needed for GNSS. An NF below 5 dB and operation without surface acoustic wave (SAW) filters under strong adjacent channel interferers are among its most important characteristics. The TX chain includes DAC, filter, mixer, and preamplification stages. Integration of the PA was demonstrated for NB-IoT in other works, like [57,58]. However, it was not considered in this project due to the additional complexity of also supporting GSM and eMTC while maintaining the high PA efficiency needed to minimize energy consumption.

The backbone of the processor system is a quad-core RISC-V cluster clocked at up to 100 MHz. The four cores host a multi-threaded

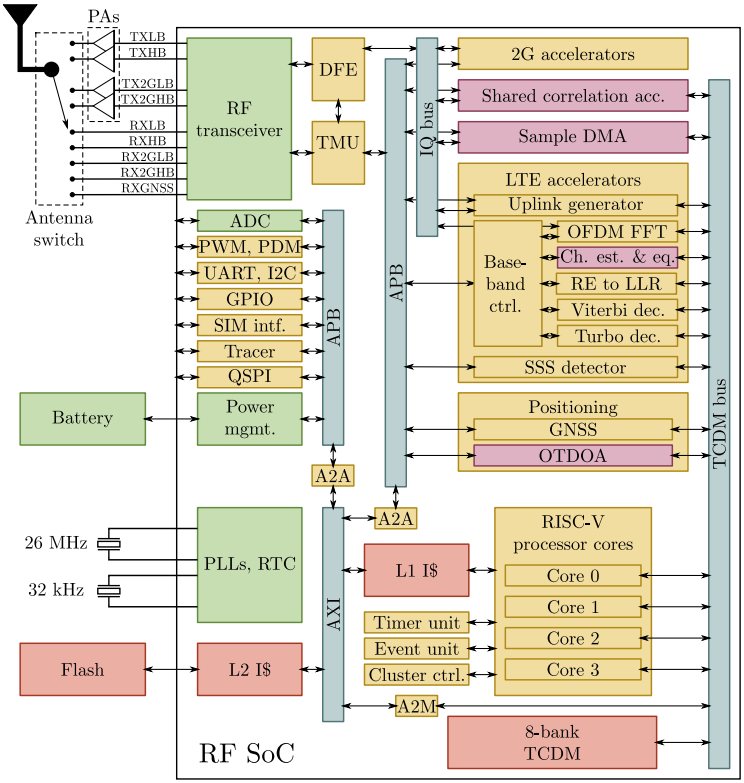

Figure 5.3: Overview of system architecture [13].

real-time operating system, the protocol stack, and part of the PHY processing. They have two levels of common instruction caches and share an 8-bank static random-access memory (SRAM) scratchpad tightly coupled data memory (TCDM). The TCDM is interfaced over a logarithmic interconnect that allows up to eight parallel accesses and automatically serializes conflicting transactions. The number of units with direct access to the TCDM interface can be higher than the number of memory banks. However, the bandwidth and latency of each link deteriorate if many of them are active at the same time.

Furthermore, adding TCDM ports increases the complexity of the interconnect and can impact the achievable clock frequency. Thus, the TCDM ports are a critical resource and should be used sparingly. The SoC communication system is built around an Advanced eXtensible Interface (AXI) [59] bus while peripherals are accessed and controlled through a simpler Advanced Peripheral Bus (APB) [60] interface and can trigger events and interrupt requests (IRQs) through dedicated interrupt lines when needed.

The DFE, at the interface between the RF transceiver and digital systems, allows digital signal scaling, filtering and DC suppression before forwarding the samples to the DBB accelerators for further processing. The RF subsystem is controlled by the time management unit (TMU), which is also responsible for the timekeeping for cellular communications and makes sure that the configuration of the transceiver is applied at the right time. For example, timing corrections and DCXO tuning are applied by the TMU, which can itself be controlled via APB and send IRQs to the processors to keep the SW synchronized to the PHY timing.

A variety of DBB accelerators are included in the system. Like most peripherals, they are configured by SW through an APB interface and can have dedicated IRQ lines. They can fetch their data from the DFE, from other accelerators, or directly from the TCDM over the low-latency interface. Direct memory access facilitates data exchange between accelerators as well as between HW and SW processing tasks. Data stored in the main processor memory can also easily be accessed from SW for debugging purposes. Furthermore, the same memory space on the TCDM can be easily shared for operations that cannot run at the same time, like initial synchronization and decoding of control channels. To fully exploit single-cycle accesses to the shared memory and to avoid the overhead generated by clock domain crossings, the accelerators that have access to the TCDM have the same clock as the processor system. The TCDM being a shared resource, conflicts in the transactions are possible, e.g. when the same memory bank is being accessed by two entities at the same time. To cope with this uncertainty, a stall probability of 20 % was assumed when assessing the performance of blocks with direct memory access.

The hardware accelerators that are relevant for the implementation of tracking for eMTC and OTDOA positioning for NB-IoT are

highlighted in pink in Fig. 5.3. The sample direct memory access (DMA) simply forwards the IQ samples from the DFE to a configurable location in the TCDM where they are buffered before further processing. The OFDM FFT also operates on the DFE output. It performs the CP removal, an FFT for the extraction, and the selection of the subcarriers that are then stored in the TCDM. The shared correlation accelerator (SCA) computes correlations in the frequency domain using the overlap and save method as described in [13]. It was designed to allow the implementation of cross-correlation-based ML PSS detection for initial synchronization, but can also correlate samples from the DFE or the TCDM with arbitrary sequences stored in the memory. It is built around a 2048-FFT that can also be used standalone on data stored in the TCDM. The channel estimation and equalization accelerator, and the OTDOA time domain correlator will be described in detail in the scope of the implementation of tracking and RSTD measurements for positioning, respectively.

5.3 Channel Estimation and Tracking

The tracking algorithm described in Chap. 3 can be realized using a combination of accelerators, originally designed for other tasks, and software for operations that do not have dedicated hardware. Before discussing the implementation of the time and frequency offset estimation, the channel estimation and equalization accelerator, also involved in tracking, is presented to highlight the importance of accurate synchronization in the decoding performance at low SNRs.

Due to the 1 ms latency between the end of an MPDCCH and its scheduled PDSCH, and to leave enough time for the other operations of the digital RX chain, a time budget of 100 µs was allocated to the eMTC channel estimation and equalization part. In that time, the channel estimation of up to four antenna ports, with at most 72 reference symbols each, must be calculated. Then, a maximum of 816 payload REs must be equalized for each AP with the final channel estimation of up to 72 subcarriers each. Assuming that one RE can be equalized per 100 MHz clock cycle, the equalization alone would require approximately 33 µs, a considerable portion of the time

budget. Thus, channel estimation and equalization cannot rely on SW and need a dedicated HW accelerator.

5.3.1 Hardware Accelerator for Channel Estimation and Equalization

The operations implemented in hardware are illustrated in Fig. 5.4 as presented in [13]. The circled numbers represent the order of the steps. First, the received REs $Y_{k,l}$ carrying reference signals on the l^{th} symbol and k^{th} subcarrier, and the expected reference symbols, $X_{k,l}$, are used to calculate the LS channel estimation $\hat{H}_{k,l}$, which is written back to the TCDM to be reused for the time and frequency offset estimation. To improve the channel estimation, the results are then combined in the time and frequency domains. Channel estimations for the same subcarrier are first averaged over one subframe and then combined with the values stored from the previous subframe with a configurable update rate α. The latter cross-subframe combination step can only be applied if the concerned subframes are coherent, i.e. if they are received consecutively and on the same narrowband. Following, the channel estimations are averaged in the frequency domain over a window of configurable size and interpolated on the subcarriers that do not carry reference symbols. The noise power is estimated from the difference of neighbour channel estimations before cross-subframe averaging. Finally, the interpolated values and the noise estimation are combined to obtain an inverted channel $H_{\text{inv}}(k)$ that is used for the minimum mean squared error (MMSE) equalization of the data symbols.

The accelerator supports equalization for channels with DMRS as well as CRS for up to four antenna ports. Besides the equalized symbols, the accelerator also generates estimations of noise and reference symbol powers, which are used for reference signal measurements and radio link monitoring.

To keep the overall latency within the defined constraints, the equalization stage must be able to process one RE per clock cycle. Since the payload symbols must be read and written back to the TCDM, at least two memory ports are needed to achieve the desired throughput. To avoid additional access to the memory interface,

the inverted channel estimations, used to equalize the data symbols, are stored in a local SRAM. The steps of the overall channel estimation and equalization procedure are serialized to allow hardware reuse. Thus, only one local data memory and two TCDM ports are needed overall. Furthermore, all complex and real multiplications were implemented with a set of 14 shared fixed-point multipliers. The fixed-point format was chosen to have a negligible impact on performance, compared to a double-precision floating-point implementation. The accelerator needs 98.8 µs to equalize MPDCCH mapped on four

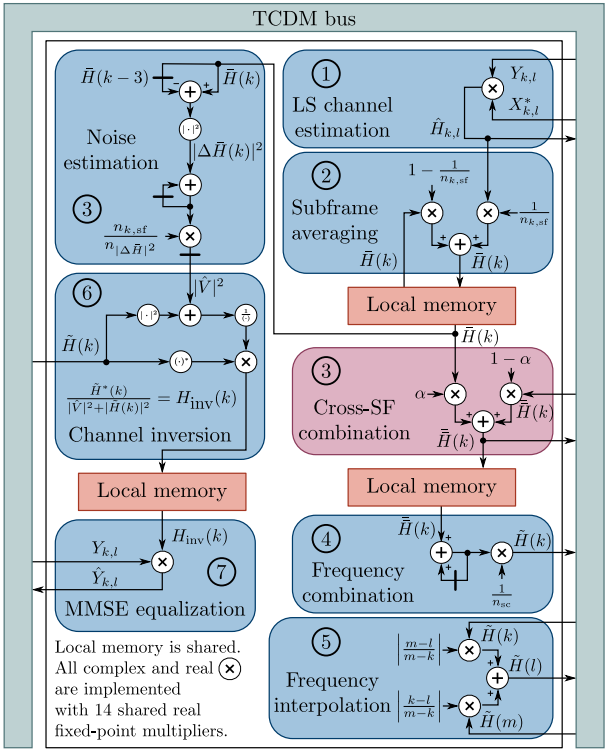

Figure 5.4: Diagram of channel estimation and equalization hardware accelerator [13].

antenna ports, closely meeting the timing constraints while keeping its hardware complexity and the impact on the overall system resources low.

The equalized data symbols are converted to LLRs by another accelerator and stored in the TCDM to serve as input for the channel decoders. The unit that prepares the LLR buffer is named RE to LLR in the system overview of Fig. 5.3. Besides QAM demodulation, deinterleaving and inverse rate matching, it also performs the accumulation of data repetitions across SFs in the LLR buffer. The coherency of the accumulated values depends mainly on the accuracy of the channel estimation. Even though time and frequency averaging help improve the channel estimation at low SNR, even for multipath fading channels, their benefits are limited by the Doppler frequency and delay spread of the channel. Furthermore, coherency across subframes and subcarriers is worsened by residual time and frequency errors, respectively. Especially, frequency offsets limit the gains of data channel repetitions and affect the achievable coverage extension.

5.3.2 Impact of Frequency Offset on Decoding Performance

To assess the impact of the frequency offset (FO) on the link budget, the PDSCH decoding performance was simulated for different configurations of the cross-SF channel estimation and different residual FOs. The corresponding results are depicted in Fig. 5.5. The lowest MCS and the highest number of repetitions in CE mode A were used for the evaluation, i.e. MCS 0 and 32 repetitions. The channel estimation was configured to operate without cross-SF combination or with cross-SF combination and an update rate of 0.5 or 0.25. The residual FO was assumed to be equally distributed with a given maximum absolute value, as for the simulations in Chap. 3. The SNR on the y-axis corresponds to the 5% BLER point which serves as the reference BLER for the sensitivity requirements of 3GPP specifications. The x-axis corresponds to the maximum frequency used to generate the residual FO.

Without frequency offset, the best performance is achieved with the smallest evaluated update rate which corresponds to a running average of 4 SFs. The performance is 0.8 dB better than with an

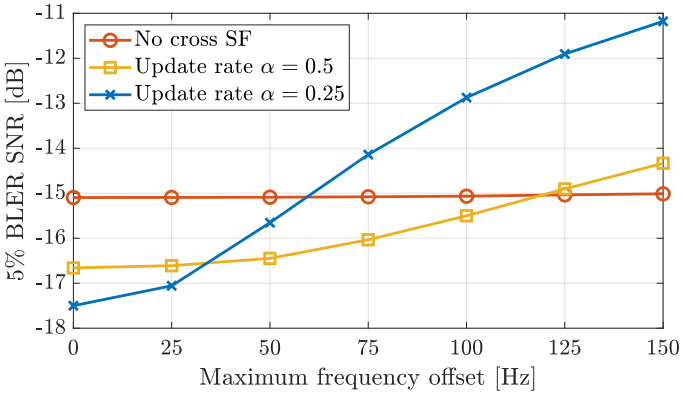

Figure 5.5: PDSCH decoding performance with MCS 0 and 32 repetitions for different configurations of the cross-subframe channel estimation and different maximum residual frequency offsets. Simulated with static AWGN channel and 10000 iterations per SNR point.

update rate of 0.5 which corresponds to a running average of 2 SFs and 2.4 dB better than simulations without the cross-SF combination. However, the smaller the update rate is, the faster the performance degrades with increasing frequency offsets. With a maximum FO of 34 Hz, the running average over 2 SFs performs already better than over 4 SFs. With an FO of 59 Hz, the performance with the running average over 4 SFs drops below the performance without cross-SF channel estimation. Therefore, accurate and robust frequency synchronization is crucial to improve channel estimation and fully exploit data repetitions in extended coverage.

5.3.3 Implementation of Time and Frequency Offset Tracking in eMTC

As discussed in Chap. 3, various approaches are possible for fine time and frequency synchronization in eMTC. Algorithms based on CRS were chosen because they provide higher accuracy and can be used

across the whole bandwidth of cells. Among these, phase difference-based methods can be easily implemented with the LS channel estimation provided by the previously presented hardware accelerator. However, the FFT-based ML approach for time and frequency offset estimation was chosen instead, despite its higher computational complexity and memory requirements, due to its superior performance in extended coverage. Even though no hard constraint on the available time for TO and FO correction is imposed by the standard, tracking must run in parallel to other procedures and can impact the load of the processor system, especially in connected mode. Furthermore, TCDM space is limited and the memory used for FO and TO estimation cannot be reused by other connected mode procedures, because tracking must run regularly together with other tasks to maintain the connection. Hence, the available resources have to be efficiently used in the implementation of fine synchronization, to avoid excessive memory usage and system load.

The different steps of the time and frequency offset estimation for tracking are summarized in the flow chart of Fig. 5.6. The configuration of the relevant hardware accelerators is controlled by a dedicated software thread that also runs part of the signal processing. Tracking is started by other tasks, e.g. after successful initial synchronization or at wake-up from a sleep period, and is periodically triggered by a timer to maintain synchronization. When a tracking event is initiated, the thread schedules a continuous reception of DL subframes including OFDM demodulation.

At the beginning of a subframe, the software configures the channel estimation accelerator to perform LS channel estimations on CRS and write them to a specific TCDM location. The channel estimation is started as soon as the OFDM demodulation of all symbols within the SF is completed. An IRQ is triggered when the channel estimation terminates. Following, the software thread combines the channel estimations in groups of two or four depending on the desired frequency estimation range, as described in Sec. 3.5.2. To reduce the computational load and energy consumption, FO estimation attempts are not performed after every received SF, but when selected SF numbers are reached. The stream of received subframes can be interrupted by frequency hops or UL transmissions if the UE is active. In that case, only bursts of consecutive subframes received on the same narrowband

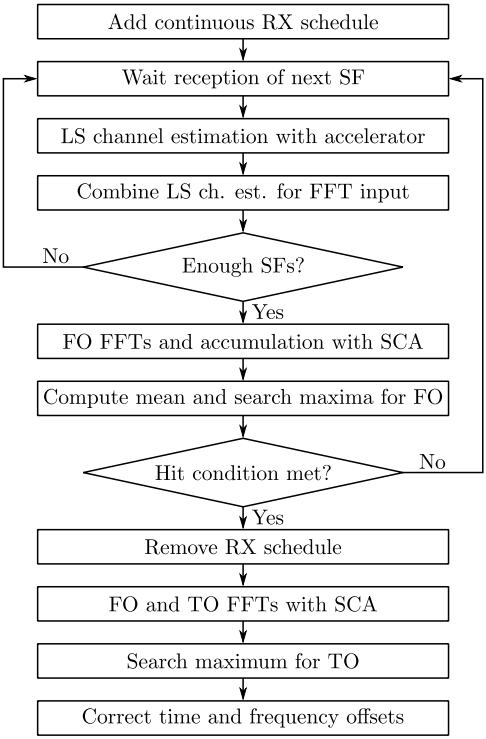

Figure 5.6: Flowchart of time and frequency offset tracking.

can be coherently accumulated within the same FFT. To accumulate the FO results of non-coherent bursts, their FFT powers are added instead of concatenating them within the same FFT, as described in Chap. 3.

If enough subframes were received, the SCA is used to calculate the 2048-FFT of the combined channel estimations, each input bin corresponding to a symbol or a slot, for the six available resource blocks. The software thread successively prepares the input buffer and configures the SCA. After each FFT, it calculates the power of the output and accumulates it with the previous results. Even though the signal processing steps implemented in software have a

small complexity compared to the computation of the FFTs and can be efficiently implemented using some of the RISC-V instruction set architecture (ISA) extensions, like the dot products used for the power calculation, they increase the usage of processor resources. Furthermore, the processing must be done in time to allow the next SF to be handled, the first deadline being the configuration of the channel estimation. Thus, an improved SCA capable of zero padding at the input and power accumulation is scheduled for the next version of the SoC.

The peak-to-second-peak and peak-to-average ratios of the accumulated FFT outputs are computed and used as stop conditions for the FO estimation. On the one hand, hard thresholds are needed for the hit criteria to prevent the correction of misestimated offsets that could break the communication link. On the other hand, more relaxed conditions allow shorter receiver-on times and lower energy consumption. The thresholds were set empirically such that no false hits are observed during the system's operation. If the stop conditions are not met, the software waits for the reception of the next SF and starts the estimation process again. If they are fulfilled, the reception of subframes for tracking is stopped and the FO is estimated from the bin index with maximum power.

The next step, TO estimation, requires the complex-valued outputs of the FO FFT of each resource block at the bin corresponding to the estimated frequency offset. At this point, the values have already been calculated with the SCA during FO estimation. However, storing the full FFT output for all resource blocks requires almost 50 kB of memory, 32 bits or 4 bytes are needed for each complex value. Since no further SFs must be received, there are not any latency constraints left for the remaining calculations. Thus, instead of storing them, the values are calculated again for each PRB, and only the bin corresponding to the estimated FO is stored for the input of the TO FFT across PRBs. The accumulation buffer of the FO FFTs can be reused as the output buffer of the SCA for all remaining FFTs, saving 50 kB of memory at the cost of additional latency and energy needed to calculate the FFTs again. The TO FFT's bin with maximum power is then searched and its index is converted into the final time offset estimation in software. Finally, the TMU is configured to correct the estimated offsets.

5.4 OTDOA Implementation[1]

OTDOA positioning requires the implementation of the LPP on the UE. The LPP is an additional layer on top of the cellular protocol stack and is also hosted by the RISC-V processor system. The LPP OTDOA positioning procedure relies on RSTD measurements that must be performed by the PHY. The most computationally intensive part of the RSTD measurement algorithm, implemented on the SoC as presented in Chap. 4, is the subframe-wise cross-correlation of the incoming samples with the expected reference signals.

To compute the correlations offline, the received samples for the whole accumulation period would have to be buffered. For NB-IoT, if the UE uses all 320 NPRS subframes that are available in standard conformance tests, almost 2.5 MB of memory would be required to store the input signal for post-processing, well beyond the memory available on the system. Thus, the UE must be able to correlate and accumulate the results on-the-fly.

Counting four real operations for each complex multiplication, the correlations need a throughput of 7.68 million real multiply-accumulate (MAC) operations per second for each time offset candidate of a single cell. If the correlations are computed for four cells and 50 time-offset candidates, this already results in over 1 GMAC/s. As this throughput could not be achieved with the embedded processor system, a dedicated hardware accelerator was designed.

5.4.1 OTDOA Hardware Accelerator

The OTDOA correlator reads the received samples from the main processor memory, correlates them with the automatically generated NPRS waveform, and writes the result back to the memory. A sketch of the dedicated correlator's architecture is shown in Fig. 5.7.

The reference signal generator reads an NPRS configuration vector for each OFDM symbol directly from the memory and generates the corresponding waveform. It takes advantage of the structure of NPRS by building the OFDM symbols as a superposition of two sinusoids modulated with a reference symbol $S_{k,l}$ each. This can be done

[1]This section was adapted from [50] published in 2020 IEEE/ION PLANS.

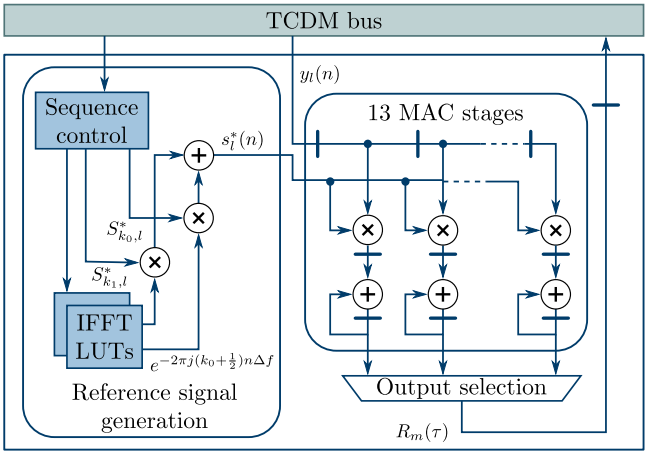

Figure 5.7: Architecture of the OTDOA hardware correlator [50].

because NPRS only occupy two out of twelve subcarriers. The sinusoids can be easily generated by cyclically reading a lookup table (LUT) of twiddle factors at equally spaced intervals. With its 13 parallel complex MAC units, the accelerator can perform up to 5.2 GMAC/s and handle up to 660 time-offset candidates in real-time. This requires only two subframes to be stored at a time, reducing the memory requirement for the sample buffer down to 16 kB which can be easily allocated in the main processor memory. The whole correlator occupies only approximately 1.6% of the overall SoC area and could be reused for neighbour cell measurements which require similar correlations.

5.4.2 HW/SW Codesign for RSTD Measurement

Unlike the cross-correlation, the other steps of the RSTD estimation algorithm have low complexity and can be handled by the RISC-V cores. Fig. 5.8 depicts the activity of different blocks during an RSTD measurement using N subframes.

The measurement starts when the connection to the base station is released and the device is in the idle state. The receiver is switched on

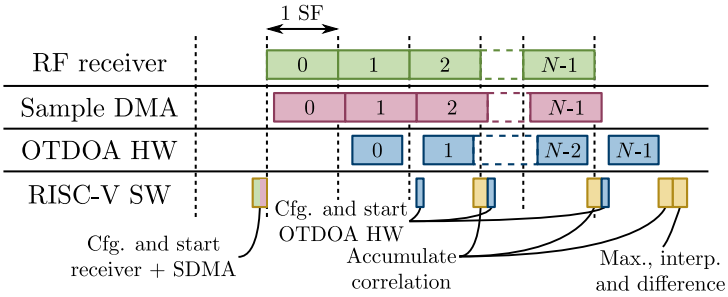

Figure 5.8: HW/SW scheduling for RSTD measurements [50].

at the beginning of the closest positioning occasion and is kept on until the last NPRS SF is received. Thus, the phase across subframes is maintained except for the drift caused by residual sampling frequency offsets. As long as the receiver is turned on the DFE outputs a stream of IQ samples at 1.92 MSamples/s. The sample DMA writes the incoming samples directly to the main processor memory where they are buffered until being processed by the OTDOA correlator. The RISC-V cores are mainly responsible for keeping track of the system's state and consequently configuring and starting the accelerators. The only online processing performed by the cores is the accumulation of the correlator's output after each subframe. For the example where the correlation is performed for four cells and 50 time offsets each the accumulation only needs approximately 400 kMAC/s, orders of magnitude less than the correlation itself. Software-based accumulation allows the developer to freely choose the combination strategy without affecting the complexity of the OTDOA accelerator. The maximum detection and the interpolation as described in (4.6) use the accumulated correlation after all SFs have been processed. The interpolation is rounded to a resolution of 30.72 MSamples/s, approximately equivalent to 10 meters. At this point a high-resolution TOA estimate for the reference and the neighbour cells is available. The last step is to subtract the TOA of the reference cell from the ones of the neighbour cells. The RSTD estimation for each neighbour cell is finally reported to higher-layer software that handles the LPP information transfers.

Chapter 6

Modem Measurements

Along with a high integration level, extended coverage and long battery life are among the most valuable features of cellular IoT (cIoT) devices. For the UE, coverage enhancement is associated with stringent requirements on the signal levels at which reliable communication and accurate measurements for positioning should be possible. Furthermore, reducing the power consumption needs the efficient use of sleep states to fully exploit the duty cycling enabled by extended discontinuous reception (eDRX) and power saving mode (PSM). Energy saving is made even more difficult when combined with extended coverage operation. The measurements presented in this chapter demonstrate the influence of the algorithmic and implementation decisions introduced previously on cellular OTDOA positioning at low input power levels and the achievable RX sensitivity of DL data channels. It is also shown how improvements in decoding performance can translate into power saving at low input signal levels.

The first section of this chapter presents the evaluation board and the measurement setup used for sensitivity and power evaluations. The second section introduces NB-IoT OTDOA performance measurements, while the third one evaluates the FO tracking performance in eMTC. In the fourth and fifth sections, the sensitivity and the power consumption of the modem presented in this work are discussed and compared with measurements performed on commercially available state-of-the-art devices.

6.1 Evaluation Board and Measurement Setup

The overall system architecture described in Chap. 5 is common to the multimodal RF SoCs presented in [56] and [13], with additional hardware accelerators included in the latter to extend the supported functionalities. The two SoCs were fabricated in a standard 110 nm complementary metal-oxide semiconductor (CMOS) technology and have an area of 32.5 mm^2 and 37.2 mm^2, respectively. The OTDOA measurements hereafter were performed on the first published SoC [56], which supports legacy GSM, EC-GSM and NB-IoT for cellular communication as well as Global Positioning System (GPS) and OTDOA positioning. While the chip depicted in Fig. 6.1 with additional eMTC communication and Beidou positioning support, as presented in [13], was used for the eMTC FO synchronization, decoding performance, and power consumption measurements discussed in this chapter.

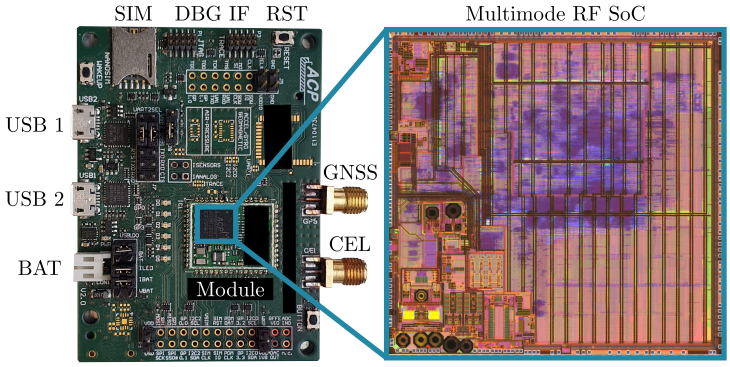

Figure 6.1: Die micrograph and evaluation board of the RF SoC presented in [13].

The evaluation board used for the software development and the different measurements is also shown in Fig. 6.1. The board, designed by a partner company, includes a modem module as well as different interfaces for debugging and standard inputs and outputs. Besides the

multimode RF SoC, the module sub-system also includes an antenna switch module (ASM), PAs, and crystal oscillators. Integrating RF, DBB processing, and CPU subsystem into a single SoC allowed the development of a compact modem module fitting within 320 mm^2. This solution supports three RATs, namely NB-IoT, eMTC and GSM while having a form factor similar to state-of-the-art dual-mode modems supporting only NB-IoT and eMTC [13].

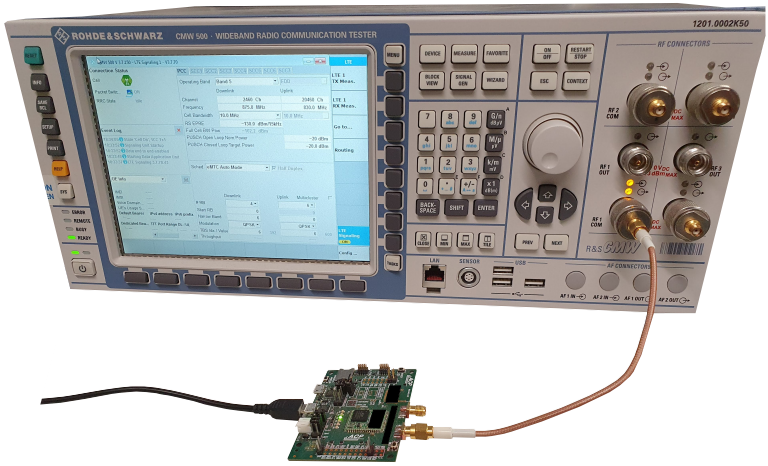

Figure 6.2: RF performance measurement set-up.

The RF measurement set-up for the performance evaluation in OTDOA positioning and cellular communication modes is shown in Fig. 6.2. The evaluation board is connected via USB to a controlling computer for UART interfacing. The measurements were done with a communication tester connected to the RF SMA connector of the board. The test equipment was operated either as a signalling unit or as an arbitrary waveform generator (ARB). In signalling mode, the communication tester emulates the behaviour of an NB-IoT or eMTC base station, including the higher protocol layers. The ARB mode, employed to generate cyclic RF signals carrying configurable baseband waveforms, provides additional flexibility to evaluate configurations that are not supported by the signalling modes.

6.2 NB-IoT OTDOA Performance[1]

The OTDOA performance was assessed by sweeping the RMS signal power at the antenna connector and evaluating the RSTD measurement hit rate using 1000 iterations for each power level. For this measurement, the test equipment was used in ARB mode with a waveform including the overlaid signals of four base stations. Constant relative delays between the signals of different cells were used for all measurements. The delays were set according to the same geometry assumed for the simulations in Chap. 4. Fig. 6.3 shows the achieved RSTD hit rate for the weakest cell using groups of $n_c \in \{1, 2, 4\}$ coherently accumulated subframes and a total of $n_{sf} = 320$ NRS SFs accumulated for each measurement. An RSTD measurement is considered a hit if it is within the equivalent of 195 m of its actual value, in which case it meets the accuracy requirement specified for normal coverage conditions listed in Tbl. 4.1.

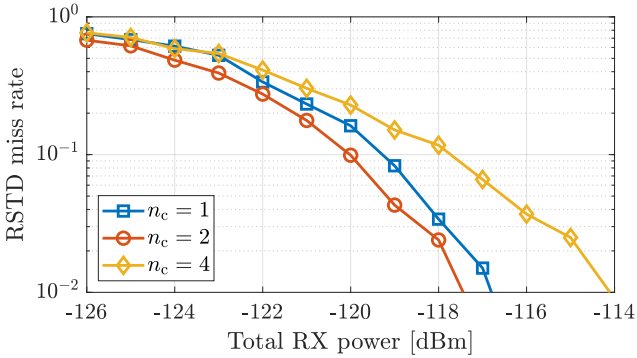

Figure 6.3: Measured RSTD miss rate for different numbers of coherently accumulated SFs [50].

The best results are achieved with $n_c = 2$ which shows approximately 0.7 dB improvement compared to $n_c = 1$, i.e. non-coherent SF combination. As already mentioned in Chap. 4, the performance

[1]This section was adapted from [50] published in 2020 IEEE/ION PLANS.

degradation for $n_c = 4$ can be caused by a residual FO. Thus, improving the FO tracking of the modem can help enhance the performance when higher coherent accumulation levels, like $n_c = 4$, are used. However, FO mismatches between the base stations, which are not emulated in the presented setup, and other non-idealities of real conditions in the field can also limit the phase coherency over periods longer than a few milliseconds.

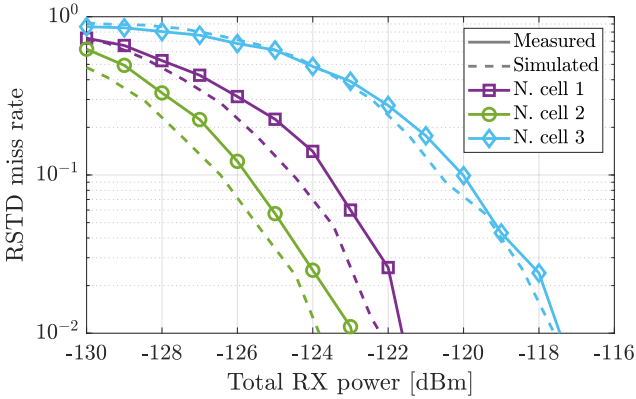

Figure 6.4: Measured RSTD miss rate for for all three neighbour cells [50].

Fig. 6.4 shows the measured RSTD hit rate for all three neighbour cells using $n_c = 2$ coherently accumulated subframes. Measurement results (solid lines) are plotted along with simulated data (dashed lines) in the same figure. The simulations were run with equally distributed frequency offsets between ± 100 Hz. The simulated FO range was chosen based on the results presented in Chap. 4 to approximately match the relative performance difference seen in the RF measurements with different values of n_c. The curves of Fig. 6.3 show better measurements with $n_c = 1$ than with $n_c = 4$, while in the simulations shown in Fig. 4.7 this was only the case for maximal FOs greater or equal to 100 Hz. The SNR values of simulation results were converted into RX power levels assuming a noise figure (NF) of 4 dB

and taking into account the power level corresponding to each cell in the final waveform used for the RF signal generation.

Thus, a total RX power of −116 dBm results in SNR levels of 0.6 dB for the reference cell, and −11 dB, −9.1 dB and −15 dB for the three neighbour cells. At this power level, the SNR of the weakest cell corresponds to the minimum specified for extended coverage operation. Results at this power level show RSTD measurement hit rates of 100% for the two stronger neighbour cells and above 99% for the weakest neighbour cell, meeting the accuracy requirement for normal coverage while being at lower SNR than specified. Thus, the specified accuracy requirements of 195 m for normal coverage and 312 m for extended coverage are both met.

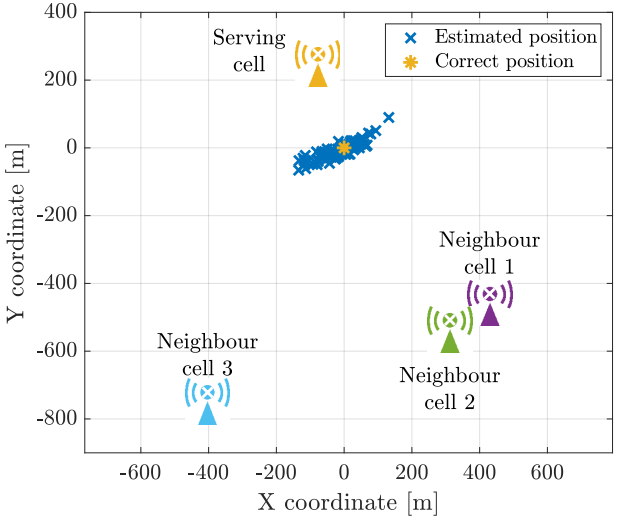

Figure 6.5: Scatter plot of position estimation at -116 dBm RX power [50].

To evaluate the expected positioning accuracy that can be achieved with the developed testbed, the RSTD measurements were exported to MATLAB and used to estimate a 2D position. The scatter plot of Fig. 6.5 depicts position estimates resulting from 100 measurements

performed at −116 dBm RX power. The RMS positioning estimation error of all 1000 measurements at this power level is 65.5 m. At −114 dBm RX power, where the SNR of the weakest cell is at the minimum requirement for normal coverage, i.e. −13 dB, the measured RMS positioning error is reduced to 48.5 m.

6.3 eMTC Tracking Performance

The padded and the packed version of the ML reduced-complexity (RC) time and frequency offset estimator using CRS presented in Sec. 3.5.2 were both implemented as described in Sec. 5.3.3. Their performance was assessed by sweeping the signal power of an LTE cell, emulating a static (ST) channel. For comparison, the phase difference (PD) frequency offset estimation with CRS, as proposed in [39], was also implemented in software using the output of the channel estimation accelerator. The FO estimation miss rates for the different variants are plotted in Fig. 6.6 against the RX power level. Again, measurements are depicted with solid lines and simulations with dashed lines. The NF of the analog receiver in eMTC measurements is 2.4 dB, the difference compared to OTDOA measurements comes from an improvement in board design. For the ML RC estimators all SFs were accumulated coherently and without early termination, using the maximum number of SFs allowed by the 2048-FFT of the SCA, i.e. 146 SFs for the padded version and 1024 SFs for the packed version. The performance of the PD algorithm was evaluated for the same numbers of accumulated SFs.

For the ML RC algorithms, thresholds for the peak-to-average and peak-to-second-peak power ratios were used as hit conditions. A miss was counted when one of the two ratios was lower than its specified threshold. The thresholds were chosen empirically and with the help of simulation results such that no false hits would be observed. The peak-to-average threshold was set to 4 and the peak-to-second-peak to 1.74. If the hit thresholds are met, the two ML RC algorithms provide a very accurate estimation of the frequency offset that can be used as a verification for the other algorithms. Thus, for the PD approach, a miss was counted when the estimated FO was more than 100 Hz away from the value estimated by the ML RC algorithm in

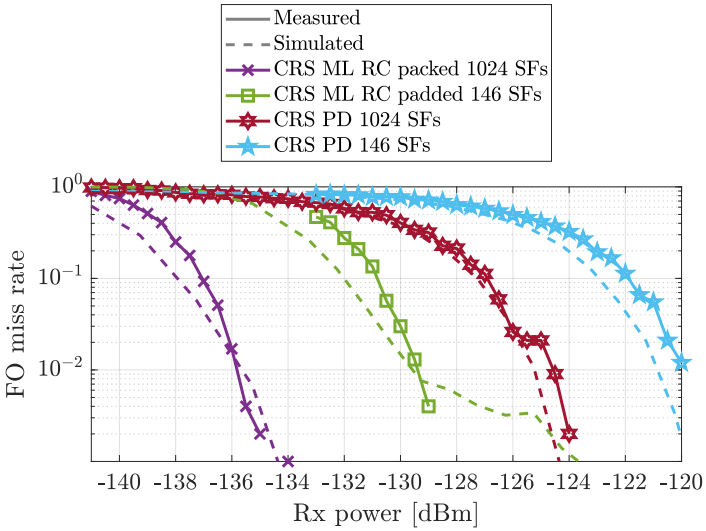

Figure 6.6: Measured (1000 iterations per SNR point) and simulated (10000 iterations per SNR point) FO estimation performance in a ST channel as partially presented in [13].

the same iteration. Iterations, where the ML RC algorithms did not fulfil the hit condition, were not counted in the evaluation of the PD algorithm.

The curves in Fig. 6.6 show comparable results between simulations and measurements. Furthermore, the measurements confirm the SNR gain of the ML RC algorithms compared to PD-based approaches commonly used for synchronization with CRS. The ML RC algorithms proposed in this work allow time and frequency tracking with at least a 10 % hit rate at RX power levels as low as −137 dBm [13]. When operated with higher input levels they also need significantly less signal accumulation to accurately estimate the time and frequency offsets. For instance, the padded ML RC algorithm with 146 SFs performs much better than the PD algorithm with 1024 SFs, almost 7 times more. Simulations show that the difference in performance would even

be larger with more stringent requirements on the frequency error, the ML RC algorithms providing more accurate estimations at low SNR. The threshold of 100 Hz was chosen to fulfil the standard requirements for UL transmissions. However, an even lower residual frequency error is advantageous to improving DL decoding performance.

6.4 Receiver Sensitivity

To extend their coverage compared to legacy GSM and LTE networks, cellular IoT (cIoT) standards target a 20 dB increase in their achievable MCL. As discussed in Chap. 2, most of the required gain must come from receiver sensitivity improvements achieved by using automatic repetitions of transmitted signals. The ability to fully exploit the available repetitions to reach the best sensitivity is a crucial feature of cIoT devices aiming to fulfil the coverage extension goal. The sensitivity of the NB-IoT and eMTC data channel was measured with different configurations for the modem presented in this work and for different commercially available devices [13]. The measured sensitivities relevant to this discussion are summarized in Tbl. 6.1. The results are compared only with the best commercial device for conciseness.

The measurements, except for eMTC coverage enhancement (CE) mode B, were performed in signalling mode with the evaluated device attached to the test equipment, the data channels scheduled

(N)PDSCH configuration	Presented in this work [13]	Commercial device [61]
eMTC 1 rep. MCS 4	−112.3 dBm	−111.3 dBm
eMTC 32 rep. MCS 0 (CE A)	−128.5 dBm	−125.0 dBm
eMTC 512 rep. MCS 0 (CE B)	−136.8 dBm	not supported
NB-IoT 1 rep. MCS 6	−117.1 dBm	−115.5 dBm
NB-IoT 512 rep. MCS 0	−137.7 dBm	−136.9 dBm

Table 6.1: Measured 5% miss rate of eMTC PDSCH and NB-IoT NPDSCH for different coverage level scenarios as presented in [13].

through the corresponding control channels, and the block error rate (BLER) determined based on the UEs' HARQ feedback. The attach procedure was carried out with a high input level, and then the power was progressively decreased to measure the BLER at different RX power levels. The CE mode B measurements, i.e. for more than 32 repetitions in eMTC, were carried out in ARB mode with periodically mapped PDSCH transmissions due to missing support by the test equipment. In both cases, the synchronization to the BS was maintained only through the UEs' time and frequency offset tracking. The sensitivity values listed in Tbl. 6.1 are defined as the input power levels at which the BLER is 5%, where only the power over the relevant RX bandwidth is considered. The measurements in signalling mode were done with 128 MPDCCH repetitions for eMTC, and 512 NPDCCH repetitions for NB-IoT. If the performance was limited by the ability of UEs to maintain the connection, the lowest power before connection loss was used as the sensitivity value.

The single repetition configurations correspond to the reference sensitivity with standard requirements of $-103\,\mathrm{dBm}$ and $-108.2\,\mathrm{dBm}$ for eMTC and NB-IoT, respectively [16]. The 3GPP specifications do not define reference sensitivity levels for any configuration with repetitions. For eMTC CE mode A, the configuration with the highest redundancy was chosen for the scope of this evaluation. Even though NB-IoT and eMTC CE mode B support up to 2048 repetitions, the increased redundancy yields diminishing returns at extremely low input levels. Thus, the sensitivity was evaluated for a maximum of 512 repetitions.

The modem presented in this work showed best-in-class sensitivity in all measured NB-IoT and eMTC modes and was the only one to support eMTC CE mode B when it was published in [13]. Especially, in eMTC CE mode A the presented modem shows the highest margin compared to the other evaluated devices, performing 3.5 dB better than the second-best. The tracking and channel estimation algorithms presented in this work and implemented on the SoC largely contribute to its better performance in the low SNR regime. This can be further investigated by considering the sensitivity for different numbers of repetitions depicted in Fig. 6.7.

The red line corresponds to measurements without the cross-SF combination of channel estimations, while the blue line was measured

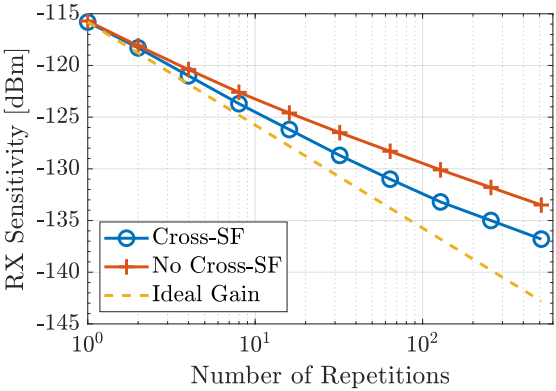

Figure 6.7: Measured eMTC sensitivity for different numbers of PDSCH repetitions with and without cross-subframe channel estimation (PDSCH, MCS 0).

with the cross-subframe combination and an update rate of $\alpha = 0.25$. The yellow dashed line represents the ideal performance based on the single-repetition case, assuming an accumulation gain of 3 dB per doubling of the number of repetitions. Cross-subframe channel estimation provides larger gains for increasing repetitions, peaking at 3.2 dB improvement for 512 repetitions. Considering the case with 32 repetitions, the sensitivity improvement is already 2.2 dB, significantly contributing to the performance margin compared to the other devices shown in Tbl. 6.1. The PDSCH decoding performance simulations presented in Sec. 5.3.2 showed that the gain of cross-SF channel estimation, with an update rate of $\alpha = 0.25$, was expected to be reduced from 2.4 dB to 2.0 dB if the maximum residual FO would increase from 0 Hz to 25 Hz, suggesting that the FO is kept below 20 Hz in the measurements of Fig. 6.7. Thus, the FO tracking solution maintains the residual frequency error in the order of ten to twenty hertz at most, despite the extremly sensitive frequency reference of the system. Even though the dashed line suggests that there is still room for improvement of several decibels for the larger numbers of repetitions, the UE would have to achieve even more accurate synchronization to

reduce the gap to the ideal achievable gain. Furthermore, at such SNR levels the performance is also limited by other tasks, like cell search and serving or neighbour cell measurements.

6.5 Power Consumption

Low energy consumption is a critical feature needed for many mMTC use cases. In cIoT systems, the high power consumption that characterizes cellular communications is compensated by allowing the UE to enter low-power modes where short activity bursts are separated by longer idle periods during which the modem can enter sleep states. To fully exploit the power-saving capabilities of cIoT standards, minimizing sleep power and modem activity are at least as important as optimizing the active power consumption of the device. The modem power consumption measured in [13] for different UE states is summarized in Tbl. 6.2. Along with the ones of the presented modem, the table also includes the results of the measured commercial UE that showed the best overall energy consumption [62]. Despite its higher power consumption when the transceiver is active, the presented device achieves similar or even better results in DRX and eDRX, when compared to other UEs. The extremely low power consumption achieved in sleep modes and the reduction of the RX-on time to the minimum required to monitor paging allow the compensation of its higher power consumption during RX. Similar results could be shown for PSM where the power of the PA during TX comes also into play for tracking area updates (TAUs). For longer inactivity periods, the power consumption in the deepest sleep state, where only the real-time clock (RTC) is active, plays an increasingly important role. With its 5.76 µW of RTC power, the presented module is well-armed to take full advantage of the long eDRX and PSM periods allowed by eMTC and NB-IoT.

Saving energy is even more challenging in extended coverage. Since repetitions are the main mechanism to increase the signal reach, CE is achieved at the cost of longer RX and TX times that drain the UE's

[3] Value from product specification [62].
[3] RX power during PDSCH subframes measured at −123 dBm RX power with 32 scheduled repetitions.

Measurement	Presented in this work [13]	Commercial device [62]
eMTC RX power	258 mW	162 mW
eMTC 18 dBm TX power	977 mW	589 mW
PSM power	5.76 µW	14.4 µW[2]
2.5 s DRX power	4.0 mW	4.8 mW
20 s eDRX power	0.5 mW	2.9 mW
eMTC CE PDSCH RX power [3]	128 mW	142 mW

Table 6.2: Measured module power in different states as presented in [13]. The power consumption was measured at the 3.6 V module power supply.

battery. The UL transmit times are mainly linked to the decoding performance of the BS and the limited TX power of mobile devices. Therefore, only limited gains can be achieved with improvements in the implementation of UEs, other than increasing the PA's efficiency. In the DL, however, the UE's ability to minimize RX-on time can help reduce power consumption. In this scope, the presented solutions for channel estimation and time and frequency offset tracking play a significant role. On the one hand, the reduced acquisition time of the RC ML FO and TO estimation algorithms presented in this work has the foreseeable advantage of allowing fine synchronization with a lower number of RX SFs compared to legacy PD algorithms. Especially, the UE always needs to restore its synchronization to the BS before monitoring paging or starting an UL transmission after long sleep cycles. In EC operation, this resynchronization step may need long accumulation times, significantly impacting the UE's power consumption. On the other hand, a better UE sensitivity does not only translate into higher coverage for the same number of repetitions but also into lower required receive times when the device is operated at higher powers.

To save power, UEs can switch off their receiver after successfully decoding a DL control or data channel. To evaluate the potential gains of such an early termination scheme, eMTC PDSCH reception in CE mode A with 32 scheduled repetitions was considered. The number of

received PDSCH subframes for successful decoding, i.e. to trigger a CRC hit, is depicted in the histograms of Fig. 6.8. The measurements were performed at −125 dBm RX power for both cases and with an update rate $\alpha = 0.25$ when the cross-subframe channel estimation was active.

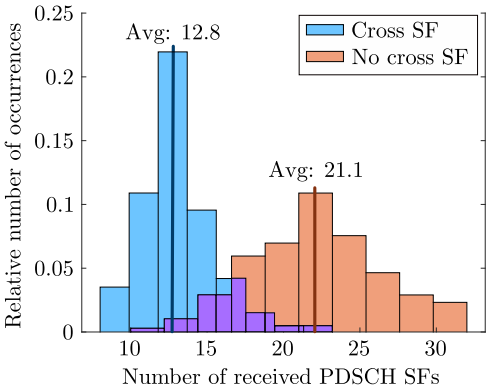

Figure 6.8: Required number of PDSCH SFs for successful decoding at −125 dBm RX power (MCS 0, 32 scheduled repetitions, 312 iterations).

Using the cross-SF combination of channel estimations helps reduce the average number of received SFs from 21.1 to 12.8. The gains in terms of power consumption during PDSCH subframes can be retrieved from the measurements presented in the first and last rows of Tbl. 6.2. While the power consumption of the presented modem during RX is almost twice as large as the one of the compared device, early termination allows a considerable reduction of the average power consumption during PDSCH subframes, bringing it slightly below the level of the commercial UE. Overall, the presented approaches for tracking and channel estimation not only improve sensitivity to reach higher coverage extension but also help reduce power consumption when combined with early termination of data reception. Additionally, more efficient usage of the available radio resources can be achieved on top of the energy savings if the BS adapts the number of scheduled repetitions to the needs of the UE.

Chapter 7

Conclusion

The emergence of low-power wide area networks (LPWAN) was expected to drive the exponential growth of MTC devices connected to the Internet. However, even though IoT connections have been steadily increasing, the evolution over the last few years did not keep up with the initial projections. The standardization efforts of the 3GPP have equipped NB-IoT and eMTC with features that should allow a massive number of connections and ubiquitous service coverage while keeping device complexity and power consumption low. But, as for other LPWAN technologies, the desired coverage enhancements and energy savings compared to legacy cellular standards are not straightforward to realize. The success of cellular IoT (cIoT) standards relies on the implementation of cost-effective modems that can fully exploit the allocated resources and take advantage of the available low-power modes.

This thesis presented solutions to extend the coverage of communication and positioning services through cellular networks and described the implementation of the proposed algorithms on low-cost integrated systems. The main contributions of this work are summarized below.

7.1 Extended Coverage Synchronization

Fine synchronization is crucial to communication and positioning systems, notably if they rely on OFDM signals. Furthermore, coverage extension support exacerbates the accuracy requirements of the local time and frequency references of cellular IoT modems while increasing the complexity of the synchronization process itself. The choice of time and frequency offset estimation algorithms impacts the achievable accuracy and needed accumulation time for the synchronization procedure of the UE. This thesis assessed different approaches for precise time and frequency offset estimation adapted to the pilot signals available in cIoT standards. The studied solutions were evaluated and compared in the scope of eMTC.

The ML-based reduced-complexity (RC) estimation developed during this work showed the best estimation accuracy and accumulation gain. The algorithm was implemented and tested on one of the modem SoCs jointly designed in this project. The proposed solution allowed the synchronization to be maintained down to $-137\,\mathrm{dBm}$ RX power, the equivalent of $-25\,\mathrm{dB}$ SNR. Compared to a legacy approach based on phase differences between reference signals, the ML RC algorithm showed a $10\,\mathrm{dB}$ performance improvement for the same accumulation period. At $-15\,\mathrm{dB}$ SNR, simulations showed that the ML RC approach could achieve the desired estimation accuracy 23.6 times faster than the legacy algorithm, considerably reducing the power needed to synchronize to a known serving cell.

7.2 Cellular OTDOA Positioning

Positioning is an essential feature needed for a variety of IoT applications. Even though location services are already widely available through GNSS, there is not a dominant solution for indoor and deep urban scenarios yet. OTDOA positioning allows reusing the cellular infrastructure to extend the reach of location services to areas not covered by GNSS signals. This work has shown that low-cost cellular modems can support OTDOA with a minimal impact on device complexity.

The evaluation of OTDOA positioning with NB-IoT showed that long accumulation periods are needed to compensate for the reduced bandwidth of the reference signals used for the time of arrival (TOA) measurements. Simulations to analyze the effect of residual frequency offsets highlighted the importance of accurate synchronization with the UE's serving cell. The presented TOA measurement strategy was implemented on a modem SoC with a dedicated hardware accelerator. Performance assessments on a laboratory set-up showed an RMS positioning error of 65.5 m at −116 dBm RX power and −15 dB SNR for the weakest cell. The achieved accuracy is within the normative requirements of the 3GPP and is sufficient for many applications.

7.3 Multimode SoCs for the Cellular IoT

The presented algorithms were evaluated on multi-mode RF SoCs supporting different communication and positioning technologies. The described SoC architecture facilitated sharing of memory and computing resources across tasks. Furthermore, the deep integration of accelerators into the processor system allowed a seamless transition between hardware and software processing for hybrid implementations.

Power and sensitivity measurements allowed the comparison of the developed SoCs with commercially available state-of-the-art devices. Despite having an active power higher than other modems, the presented solution achieves a lower average power consumption in eDRX and PSM through optimized sleep states and minimized RX time. In eMTC CE mode A, the developed SoC outperforms the best competitor by 3.5 dB in PDSCH sensitivity with 32 repetitions. In this configuration, cross-subframe channel estimation combined with accurate frequency offset tracking yields a 2.2 dB gain out of the 2.4 dB that would be possible with perfect synchronization. With CE mode B, the presented modem reaches an overall eMTC sensitivity of −136.8 dBm, 11.8 dB better than the compared devices which do not support this mode. At a fixed input power level, the sensitivity advantage translates into reduced power consumption and more efficient usage of the radio resources due to the smaller number of data repetitions needed for successful decoding.

7.4 Outlook

With appropriate UE enhancements, coverage extension through blind repetitions allows the compensation of sensitivity losses caused by the reduced number of receive antennas of low-cost MTC modems and the expansion of cellular signals' reach to remote or buried devices. However, coverage enhancement features can only be exploited if operators upgrade their infrastructure to support cellular IoT standards. In places where LTE is already available, the networks can be updated to support NB-IoT and eMTC with relatively low overhead. If new infrastructure is required, the deployment cost may be prohibitive for network operators. Therefore, communication over non-terrestrial networks with 5G technology has gained interest in recent years to use satellites and unmanned aerial vehicles (UAVs) to cover areas that cannot be reached by conventional deployments [63].

For cIoT UEs, NB-IoT and eMTC can support non-terrestrial communications with a relatively low standardization effort [64, 65]. The main challenges for the PHY lie in handling the increased propagation delay, Doppler shift, and path loss suffered by the transmitted signals. The time and frequency offset estimation methods presented in this thesis can serve as a starting point for the synchronization strategies needed to deal with the high Doppler frequency variations and the tight link budget of the different scenarios covered by satellite communications.

Appendix A

Derivations

A.1 Effect of Time and Frequency Offsets

This derivation follows the same steps as in [35], considering the effect of time offset instead of sampling frequency offset. The starting point is (3.1) from Chap. 2 for the transmitted OFDM signal in the baseband:

$$s_l(t) = \sum_{k=-\frac{N_{\mathrm{FFT}}}{2}}^{\frac{N_{\mathrm{FFT}}}{2}-1} X_{k,l} e^{j2\pi f_\delta k(t-T_{\mathrm{CP}}-lT_{\mathrm{SB}})} \ ,$$

where $lT_{\mathrm{SB}} \leq t < (l+1)T_{\mathrm{SB}}$, and $s_l(t) = 0$ otherwise. After upconversion to the carrier frequency, transmission through the radio channel, and downconversion back to baseband with a frequency offset ν, the received signal is:

$$r_l(t) = e^{j2\pi\nu t} s_l(t) * h_l(t) + n(t) \ .$$

After sampling with a sampling period T and a time offset τ and removing the cyclic prefix the signal can be represented by:

$$\begin{aligned} y_l(t) = & \sum_n \sum_k X_{k,l} e^{-j2\pi f_\delta k(lT_{\mathrm{SB}}+T_{\mathrm{CP}})} e^{j2\pi\nu t} \\ & \cdot \delta(t - nT - lT_{SB} - \tau) \left(e^{j2\pi k f_\delta t} * h_l(t)\right) \end{aligned}$$

$$+\sum_{n}\delta(t-nT-lT_{SB}-\tau)n(t)$$

where k runs from $-\frac{N_{\mathrm{FFT}}}{2}$ to $\frac{N_{\mathrm{FFT}}}{2}-1$ and n from 0 to $N_{\mathrm{FFT}}-1$.

To approximate the FFT for OFDM demodulation performed at the receiver the signal is first shifted by $lT_{\mathrm{SB}}+T_{\mathrm{CP}}+\tau$ and transformed into the Fourier domain:

$$\begin{aligned}
Y_l(f) =& \sum_{n}\sum_{k} X_{k,l}H(kf_\delta)e^{-j2\pi f_\delta k(lT_{\mathrm{SB}}+T_{\mathrm{CP}})}e^{j2\pi f(lT_{\mathrm{SB}}+T_{\mathrm{CP}}+\tau)} \\
&\cdot e^{j2\pi(nT+lT_{\mathrm{SB}}+\tau)(kf_\delta-(f-\nu))} + N_l(f) \\
=& \sum_{k} X_{k,l}H(kf_\delta)e^{-j2\pi f_\delta k(lT_{\mathrm{SB}}+T_{\mathrm{CP}})}e^{j2\pi f(lT_{\mathrm{SB}}+T_{\mathrm{CP}}+\tau)} \\
&\cdot e^{-j2\pi(lT_{\mathrm{SB}}+\tau)(f-kf_\delta-\nu)}\sum_{n} e^{-j2\pi nT(f-kf_\delta-\nu)} + N_l(f)
\end{aligned}$$

Using the formula $\sum_{n=0}^{N-1} e^{j2\alpha n} = e^{j\alpha(N-1)}\frac{\sin(\alpha N)}{\sin(\alpha)}$, the summation over the sample index n can be written in closed form:

$$\begin{aligned}
Y_l(f) =& \sum_{k} X_{k,l}H(kf_\delta)e^{-j2\pi f_\delta k(lT_{\mathrm{SB}}+T_{\mathrm{CP}})}e^{j2\pi f(lT_{\mathrm{SB}}+T_{\mathrm{CP}}+\tau)} \\
&\cdot e^{-j2\pi(lT_{\mathrm{SB}}+\tau)(f-kf_\delta-\nu)}e^{-j\pi T(f-kf_\delta-\nu)(N_{\mathrm{FFT}}-1)} \\
&\cdot \frac{\sin(\pi T(f-kf_\delta-\nu)N_{\mathrm{FFT}})}{\sin(\pi T(f-kf_\delta-\nu))} + N_l(f)
\end{aligned}$$

An estimation of the demodulated symbol is obtained by evaluating the Fourier transform at the k^{th} subcarrier $Y_{k,l} = Y_l(f)|_{f=kf_\delta}$:

$$\begin{aligned}
Y_{k,l} =& \frac{1}{N_{\mathrm{FFT}}}\sum_{\kappa} X_{\kappa,l}H(\kappa f_\delta)e^{j2\pi f_\delta(k-\kappa)(lT_{\mathrm{SB}}+T_{\mathrm{CP}})}e^{j2\pi kf_\delta\tau} \\
&\cdot e^{-j2\pi(lT_{\mathrm{SB}}+\tau)(f_\delta(k-\kappa)-\nu)}e^{-j\pi T(f_\delta(k-\kappa)-\nu)(N_{\mathrm{FFT}}-1)} \\
&\cdot \frac{\sin(\pi T(f_\delta(k-\kappa)-\nu)N_{\mathrm{FFT}})}{\sin(\pi T(f_\delta(k-\kappa)-\nu))} + N_l(kf_\delta)
\end{aligned}$$

Using $Tf_\delta = \frac{1}{N_{\mathrm{FFT}}}$ shows that with zero frequency offset $\nu = 0$ all terms with $k \neq \kappa$ disappear:

$$\frac{\sin(\pi T f_\delta (k-\kappa) N_{\mathrm{FFT}})}{\sin(\pi T f_\delta (k-\kappa))} = \frac{\sin(\pi(k-\kappa))}{\sin(\frac{\pi(k-\kappa)}{N_{\mathrm{FFT}}})} = \begin{cases} N_{\mathrm{FFT}} & \text{if } k = \kappa \\ 0 & \text{otherwise.} \end{cases}$$

Separating the desired subcarrier from the rest and approximating $\frac{\sin(\pi T \nu N_{\mathrm{FFT}})}{\sin(\pi T \nu)} \approx N_{\mathrm{FFT}}$ for small frequency offsets ν yields:

$$\begin{aligned} Y_{k,l} =& X_{k,l} H(k f_\delta) e^{j2\pi(k f_\delta \tau + \nu\tau + \nu l T_{\mathrm{SB}} + \frac{1}{2} T\nu(N_{\mathrm{FFT}}-1))} \\ &+ \frac{1}{N_{\mathrm{FFT}}} \sum_{\kappa \neq k} X_{\kappa,l} H(\kappa f_\delta) e^{j2\pi f_\delta (k-\kappa)(l T_{\mathrm{SB}} + T_{\mathrm{CP}})} e^{j2\pi k f_\delta \tau} \\ &\cdot e^{-j2\pi(l T_{\mathrm{SB}} + \tau)(f_\delta(k-\kappa)-\nu)} e^{-j\pi T(f_\delta(k-\kappa)-\nu)(N_{\mathrm{FFT}}-1)} \\ &\cdot \frac{\sin(\pi T(f_\delta(k-\kappa)-\nu) N_{\mathrm{FFT}})}{\sin(\pi T(f_\delta(k-\kappa)-\nu))} + N_l(k f_\delta) \end{aligned}$$

Finally, the constant terms of the complex phase can be included in an equivalent channel impulse response to get the final form:

$$Y_{k,l} = X_{k,l} H_{k,l} e^{j\phi(k,l,\nu,\tau)} + W(k,l,\nu,\tau) + N_{k,l} \tag{A.1}$$

where:

$$\begin{aligned} \phi(k,l,\nu,\tau) &= 2\pi(k f_\delta \tau + \nu l T_{\mathrm{SB}}) \text{ , and} \\ W(k,l,\nu,\tau) &= \frac{1}{N_{\mathrm{FFT}}} \sum_{\kappa \neq k} X_{\kappa,l} H(\kappa f_\delta) e^{j2\pi f_\delta (k-\kappa)(l T_{\mathrm{SB}} + T_{\mathrm{CP}})} e^{j2\pi k f_\delta \tau} \\ &\quad \cdot e^{-j2\pi(l T_{\mathrm{SB}} + \tau)(f_\delta(k-\kappa)-\nu)} e^{-j\pi T(f_\delta(k-\kappa)-\nu)(N_{\mathrm{FFT}}-1)} \\ &\quad \cdot \frac{\sin(\pi T(f_\delta(k-\kappa)-\nu) N_{\mathrm{FFT}})}{\sin(\pi T(f_\delta(k-\kappa)-\nu))} \text{ .} \end{aligned}$$

A.2 Cramer-Rao Lower Bound

To derive the CRLB for the joint estimation of the time and frequency offset when the transmitted symbols are known, the ICI term $W(k,l,\nu,\tau)$ in (A.1) is neglected:

$$Y_{k,l} = X_{k,l} H_{k,l} e^{j\phi(k,l,\nu,\tau)} + N_{k,l} \ .$$

Assuming a single tap constant channel $H_{k,l} = Ae^{j\alpha}$ and QPSK modulated symbols the signal after matched reception is:

$$S(k,l;A,\alpha,\nu,\tau) = Ae^{j(2\pi(kf_\delta\tau+\nu\tau+\nu l T_{\mathrm{SB}}+\frac{1}{2}T\nu(N_{\mathrm{FFT}}-1)+\alpha)} + N_{k,l} \ .$$

In this form time and frequency estimation can be handled as the parameter estimation of a two-dimensional sinusoid. Defining the parameter vector $\boldsymbol{\theta} = [A,\alpha,\nu,\tau]^{\mathrm{T}}$ and assuming IID normally distributed complex noise $N_{k,l} \sim \mathcal{CN}(0,\sigma_N^2)$, the elements of the Fisher information matrix $\boldsymbol{J}$ are [38]:

$$J_{i,j} = \frac{2}{\sigma^2} \sum_{k,l} \left[\frac{\partial S_I}{\partial \theta_i} \frac{\partial S_I}{\partial \theta_j} + \frac{\partial S_Q}{\partial \theta_i} \frac{\partial S_Q}{\partial \theta_j} \right] \ . \tag{A.2}$$

The factor two difference compared to [38] is due to a different noise definition. $S_I = S_I(k,l;\boldsymbol{\theta})$ and $S_Q = S_Q(k,l;\boldsymbol{\theta})$ represent the real and imaginary parts of the symbols, respectively. The real part is given by:

$$S_I(k,l;\boldsymbol{\theta}) = A\cos\left(2\pi(kf_\delta\tau+\nu\tau+\nu l T_{\mathrm{SB}}+\frac{1}{2}T\nu(N_{\mathrm{FFT}}-1)+\alpha\right) \ ,$$

and its partial derivatives are:

$$\begin{aligned}
\frac{\partial S_I}{\partial A} &= \cos(\phi(k,l;\boldsymbol{\theta})) \\
\frac{\partial S_I}{\partial \alpha} &= -A\sin(\phi(k,l;\boldsymbol{\theta})) \\
\frac{\partial S_I}{\partial \nu} &= -2\pi A(\tau + lT_{\mathrm{SB}} + \frac{1}{2}T(N_{\mathrm{FFT}}-1))\sin(\phi(k,l;\boldsymbol{\theta})) \\
\frac{\partial S_I}{\partial \tau} &= -2\pi A(kf_\delta+\nu)\sin(\phi(k,l;\boldsymbol{\theta}))
\end{aligned}$$

After similarly calculating the partial derivatives of $S_Q(k,l;\boldsymbol{\theta})$, the elements of the Fisher information matrix can be obtained as in [37]:

$$
\begin{aligned}
J_{1,1} &= \frac{2KL}{\sigma^2} \\
J_{1,2} &= 0 \\
J_{1,3} &= 0 \\
J_{1,4} &= 0 \\
J_{2,2} &= \frac{2A^2KL}{\sigma^2} \\
J_{2,3} &\approx \frac{2\pi A^2 T_{\mathrm{SB}} KL(L-1)}{\sigma^2} \\
J_{2,4} &\approx 0 \\
J_{3,3} &\approx \frac{4\pi^2 A^2 T_{\mathrm{SB}}^2 KL(L-1)(2L-1)}{3\sigma^2} \\
J_{3,4} &\approx 0 \\
J_{4,4} &\approx \frac{4\pi^2 A^2 f_\delta^2 LK(K+2)(K-1)}{3\sigma^2} \,.
\end{aligned}
$$

The approximations assume $L, K \gg 1$, $\nu \ll f_\delta$ and $\tau \ll T_{\mathrm{SB}}$. The CRLBs of $\boldsymbol{\theta}$ are the diagonal elements of $\boldsymbol{J}^{-1}$:

$$
\begin{aligned}
\mathrm{CRLB}_A &= \frac{\sigma^2}{2KL} \\
\mathrm{CRLB}_\alpha &= \frac{2\sigma^2(2L-1)}{A^2KL(L+1)} \approx \frac{4\sigma^2}{A^2KL} \\
\mathrm{CRLB}_\nu &= \frac{3\sigma^2}{2\pi^2 A^2 T_{\mathrm{SB}}^2 KL(L-1)(L+1)} \approx \frac{3\sigma^2}{2\pi^2 A^2 T_{\mathrm{SB}}^2 KL^3} \\
\mathrm{CRLB}_\tau &= \frac{3\sigma^2}{2\pi^2 A^2 f_\delta^2 LK(K-1)(K-2)} \approx \frac{3\sigma^2}{2\pi^2 A^2 f_\delta^2 LK^3}
\end{aligned}
$$

A.3 Outlier Probability in Additive White Gaussian Noise

The miss rate of the estimator presented in Sec. 3.5.3 can be approximated with the outlier probability of the FO estimation metric

following a procedure similar to [38]. The ML RC algorithm uses the DFT across accumulated LS channel estimations as described by (3.21) and (3.23):

$$\tilde{H}_{m,n} = \sum_{(k,l)\in\mathcal{C}_{m,n}} \hat{H}_{k,l} \ , \tag{A.3}$$

$$\tilde{H}_m^{(p)}(v) = \mathcal{F}_{N_n}\{\tilde{H}_{m,pN+n}\} = \sum_{n=0}^{N_n-1} \tilde{H}_{m,pN+n} e^{-2\pi j \frac{nv}{N_n}} \ . \tag{A.4}$$

Neglecting the effect of the time offset and the radio channel, the LS channel estimations form a complex sinusoid in AWGN:

$$\hat{H}_{k,l} = e^{2\pi j l T_{\mathrm{SB}} \nu} + N_{k,l} \tag{A.5}$$

$N_{k,l}$ is complex Gaussian noise $N_{k,l} \sim \mathcal{CN}(0, \sigma^2)$. With the assumptions made in Sec. 3.5.3 and listed below, $\tilde{H}_m^{(p)}(v)$ can be easily characterized.

- The combination of channel estimations into $\tilde{H}_{m,n}$ described by (A.3) is lossless.
- No zero padding is needed at the input of the DFT used to calculate $\tilde{H}_m^{(p)}(v)$.
- A DFT of the same size N_n as the input vector $\tilde{H}_{m,pN+n}$ is used.
- The FO ν corresponds exactly to the frequency of one DFT output bin $\tilde{H}_m^{(p)}(v)$ in (A.4).

In this case, the signal bin $\tilde{H}_m^{(p)}(v)|_{v=s}$ has a mean N_c and variance of $N_\mathrm{c}\sigma^2$ and the noise bins $\tilde{H}_m^{(p)}(v)|_{v=s}$ have zero mean and variance of $N_\mathrm{c}\sigma^2$. N_c corresponds to the number of LS channel estimations accumulated through the combination of (A.3) and (A.4).

$$\tilde{H}_m^{(p)}(s) \sim \mathcal{CN}(N_\mathrm{c}\, N_\mathrm{c}\sigma^2)$$
$$\tilde{H}_m^{(p)}(r) \sim \mathcal{CN}(0, N_\mathrm{c}\sigma^2)$$

Using a scaled version of (3.24) yields an FO estimation metric that can be characterized by central and non-central chi-squared distributions which correspond to the addition of the power of gaussian distributions with unit variance [66]:

$$C'_\nu(v) = \frac{2}{N_\mathrm{c}\sigma^2} \sum_{p=0}^{P-1} \sum_{m=0}^{M-1} |\tilde{H}_m^{(p)}(v)|^2 \ . \tag{A.6}$$

P is the number of non-coherently accumulated time-domain blocks and M is the number of combined physical resource blocks (PRBs) in the frequency domain. The factor of two in the scaling constant comes from the definition of the complex noise variance where the real and imaginary components have $\frac{\sigma^2}{2}$ variance each.

The PDFs of the estimation metric correspond to the addition of the power of $2PM$ gaussian random variables, including the real and imaginary parts of $\tilde{H}_m^{(p)}$. Denoting by $f_s(x)$ the PDF of the signal bin $C'_\nu(s)$ and by $f_r(x)$ the PDF of the noise bins $C'_\nu(r)$ the PDFs for $x > 0$ are [66]:

$$f_s(x) = \frac{1}{2}\left(\frac{x}{\lambda}\right)^{\frac{\kappa-2}{4}} e^{-\frac{1}{2}(x+\lambda)} I_{\frac{\kappa}{2}-1}\left(\sqrt{\lambda x}\right) \ .$$
$$f_r(x) = \frac{1}{2^{\frac{\kappa}{2}}\Gamma\left(\frac{\kappa}{2}\right)} x^{\frac{\kappa}{2}-1} e^{-\frac{x}{2}} \ .$$

The parameters of the distributions are $\kappa = 2PM$ degrees of freedom and a non-centrality parameter $\lambda = \frac{2N_\mathrm{c}PM}{\sigma^2}$. $I_a(z)$ is the modified Bessel function of the first kind and order a and $\Gamma(z)$ is the Gamma function.

The outlier probability q of the FO estimation can then be calculated numerically with its integral form [38]:

$$q = 1 - \int_0^\infty f_s(x) \left[\int_0^x f_r(y)dy\right]^{N_n-1} dx \ . \tag{A.7}$$

Appendix B

Acronyms

3GPP	3rd Generation Partnership Project
4G	fourth-generation
5G	fifth-generation
AC	autocorrelation
ADC	analog-to-digital converter
A-GNSS	assisted GNSS
AP	antenna port
APB	Advanced Peripheral Bus
ARB	arbitrary waveform generator
ASM	antenna switch module
AWGN	additive white Gaussian noise
AXI	Advanced eXtensible Interface
BLE	Bluetooth Low Energy
BLER	block error rate
BS	base station
BW	bandwidth
CC	cross-correlation

CE	coverage enhancement
CFO	carrier frequency offset
cIoT	cellular IoT
CMOS	complementary metal-oxide semiconductor
CP	cyclic prefix
CPU	central processing unit
CRC	cyclic redundancy check
CRLB	Cramer-Rao lower bound
CRS	cell-specific reference signals
DAC	digital-to-analog converter
DBB	digital baseband
DC	direct current
DCI	downlink control information
DCXO	digitally-controlled crystal oscillator
DFE	digital front-end
DFT	discrete Fourier transform
DL	downlink
DMA	direct memory access
DMRS	demodulation reference signals
DRX	discontinuous reception
EC	extended coverage
E-CID	enhanced cell ID
eDRX	extended discontinuous reception
eMBB	enhanced mobile broadband
eMTC	enhanced Machine-Type Communication
EPA	extended pedestrian A
ETU	extended typical urban
EVA	extended vehicular A
FD	frequency domain
FDD	frequency-division duplex
FEC	forward error correction

FFT	fast Fourier transform
FO	frequency offset
GNSS	Global Navigation Satellite System
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HARQ	hybrid automatic repeat request
HD-FDD	half-duplex FDD
HW	hardware
ICI	inter-carrier interference
ID	identity
IFFT	inverse fast Fourier transform
IID	independent and identically distributed
IIS	Integrated Systems Laboratory
IoT	Internet of things
IRQ	interrupt request
ISA	instruction set architecture
ISI	inter-symbol interference
ITU	International Telecommunication Union
LBS	location-based services
LLR	log-likelihood ratio
LNA	low-noise amplifier
LPP	LTE Positioning Protocol
LPWAN	low-power wide area networks
LS	least squares
LTE	Long Term Evolution
LUT	lookup table
MAC	multiply-accumulate
MCL	maximum coupling loss
MCS	modulation and coding scheme

MIB	master information block
MIMO	multiple-input multiple-output
mIoT	massive IoT
ML	maximum likelihood
MMSE	minimum mean squared error
mMTC	massive machine-type communication
MPDCCH	MTC physical downlink control channel
MSE	mean squared error
MTC	machine-type communications
NB	narrowband
NB-IoT	Narrowband Internet-of-Things
NC	normal coverage
NF	noise figure
NPBCH	narrowband PBCH
NPDCCH	narrowband PDCCH
NPDSCH	narrowband PDSCH
NPRS	narrowband PRS
NPSS	narrowband PSS
NPUSCH	narrowband PUSCH
NR	New Radio
NRS	narrowband reference signals
NSSS	narrowband SSS
OFDM	orthogonal frequency-division multiplexing
OTDOA	observed time difference of arrival
PA	power amplifier
PBCH	physical broadcast channel
PCID	physical cell identity
PD	phase difference
PDCCH	physical downlink control channel
PDF	probability density function
PDSCH	physical downlink shared channel

PHY	physical layer
PRB	physical resource block
PRS	positioning reference signals
PSD	power spectral density
PSK	phase-shift keying
PSM	power saving mode
PSS	primary synchronization signal
PUCCH	physical uplink control channel
PUSCH	physical uplink shared channel
QAM	quadrature amplitude modulation
QPSK	quadrature phase-shift keying
RAT	radio access technology
RC	reduced-complexity
RE	resource element
RF	radio frequency
RMS	root-mean-square
RMSE	RMS error
RSTD	reference signal time difference
RTC	real-time clock
RX	receive
SAW	surface acoustic wave
SC-FDMA	single-carrier frequency-division multiple access
SCA	shared correlation accelerator
SDR	software-defined radio
SF	subframe
SFN	system frame number
SFO	sampling frequency offset
SINR	signal-to-interference-plus-noise ratio
SNR	signal-to-noise ratio
SoC	system-on-chip
SRAM	static random-access memory

SSS	secondary synchronization signal
ST	static
SW	software
TA	timing advance
TAU	tracking area update
TCDM	tightly coupled data memory
TCXO	temperature compensated crystal oscillator
TD	time domain
TDD	time-division duplex
TDOA	time difference of arrival
TMU	time management unit
TO	time offset
TOA	time of arrival
TTI	transmission time interval
TX	transmit
UAV	unmanned aerial vehicle
UE	user equipment
UL	uplink
URLLC	ultra-reliable and low-latency communication
WLAN	wireless local area network

Bibliography

[1] Ericsson. (2022, Jun.) Ericsson Mobility Report. [Online]. Available: https://www.ericsson.com/en/reports-and-papers/mobility-report/reports/june-2022

[2] R. S. of ITU (ITU-R), *Minimum requirements related to technical performance for IMT-2020 radio intarface(s)*, ITU Report ITU-R M.2410-0, Nov. 2017. [Online]. Available: https://www.itu.int/pub/R-REP-M.2410

[3] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The next generation wireless access technology.* Academic Press, 2020.

[4] R. S. of ITU (ITU-R), *Detailed specifications of the terrestrial radio interfaces of International Mobile Telecommunications-2020 (IMT-2020)*, ITU Recommendation ITU-R M.2150-0, Feb. 2021. [Online]. Available: https://www.itu.int/rec/R-REC-M.2150

[5] O. Liberg, M. Sundberg, E. Wang, J. Bergman, and J. Sachs, *Cellular Internet of things: technologies, standards, and performance.* Academic Press, 2017.

[6] F. Van Diggelen, *A-GPS: Assisted GPS, GNSS, and SBAS*, 2009.

[7] S. Fisher, "Observed time difference of arrival (OTDOA) positioning in 3GPP LTE," Quallcomm White Paper, Qualcomm Technologies Inc., Tech. Rep., Jun. 2014.

[8] Forbes. (2022) The Dusk Before The 3G Sunset. [Online]. Available: https://www.forbes.com/sites/forbestechcouncil/2022/05/11/the-dusk-before-the-3g-sunset

[9] Swisscom. (2021, Apr.) The 2G network is now part of Swisscom's history. [Online]. Available: https://www.swisscom.ch/en/about/news/2021/04/14-das-2g-netz-ist-bei-swisscom-geschichte.html

[10] GSMA. (2022, Feb.) Mobile IoT Deployment Map. [Online]. Available: https://www.gsma.com/iot/deployment-map/

[11] H. Kröll, "An Evolved EDGE System-on-Chip for the Cellular Internet of Things," Ph.D. dissertation, ETH Zurich, 2016.

[12] B. Weber, "An Extended Coverage System on Chip for the Cellular Internet of Things," Ph.D. dissertation, ETH Zurich, 2018.

[13] S. Lippuner, M. Salomon, M. Korb, M. Gautschi, T. Dellsperger, S. Altorfer, J. Rogin, S. Willi, D. Tschopp, B. Weber, and Q. Huang, "A triple-mode cellular IoT SoC achieving -136.8-dbm eMTC sensitivity," *IEEE Solid-State Circuits Letters*, vol. 3, pp. 418–421, Sep. 2020.

[14] S. Lippuner, "Extended Coverage Cellular Communication," Ph.D. dissertation, ETH Zurich, 2022, to be published.

[15] M. Elsaadany, A. Ali, and W. Hamouda, "Cellular LTE-A technologies for the future Internet-of-Things: Physical layer features and challenges," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2544–2572, Jul. 2017.

[16] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception*, 3GPP TS 36.101, Rev. 15.16.0, Nov. 2021. [Online]. Available: https://www.3gpp.org/DynaReport/36101.htm

[17] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[18] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation*, 3GPP TS 36.211, Rev. 14.14.0, Apr. 2020. [Online]. Available: https://www.3gpp.org/DynaReport/36211.htm

[19] E. Dahlman, S. Parkvall, and J. Skold, *4G, LTE-advanced Pro and the Road to 5G.* Academic Press, 2016.

[20] V. Saxena, J. Bergman, Y. Blankenship, A. Wallen, and H. S. Razaghi, "Reducing the modem complexity and achieving deep coverage in LTE for machine-type communications," in *2016 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7.

[21] Y.-P. E. Wang, X. Lin, A. Adhikary, A. Grovlen, Y. Sui, Y. Blankenship, J. Bergman, and H. S. Razaghi, "A primer on 3GPP narrowband internet of things," *IEEE Communications Magazine*, vol. 55, no. 3, pp. 117–123, Mar. 2017.

[22] Sierra Wireless et al. (2017, Jan.) Coverage analysis of LTE-M. [Online]. Available: https://altair-semi.com/wp-content/uploads/2017/02/Coverage-Analysis-of-LTE-CAT-M1-White-Paper.pdf

[23] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA); Base Station (BS) radio transmission and reception*, 3GPP TS 36.104, Rev. 14.9.0, May 2019. [Online]. Available: https://www.3gpp.org/DynaReport/36104.htm

[24] W. Yang, M. Wang, J. Zhang, J. Zou, M. Hua, T. Xia, and X. You, "Narrowband wireless access for low-power massive internet of things: A bandwidth perspective," *IEEE Wireless Communications*, vol. 24, no. 3, pp. 138–145, Jun. 2017.

[25] R. Faragher and R. Harle, "Location fingerprinting with Bluetooth Low Energy beacons," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2418–2428, Nov. 2015.

[26] M. Korb, Q. Huang, P. Stockel, G. C. Kappen, B. Weber, and M. Garcia, "A cellular-modem-hosted low-cost single-shot dual-mode assisted-GNSS receiver for the internet of things," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 1273–1279.

[27] R. Zekavat and R. M. Buehrer, *Handbook of position location: Theory, practice and advances.* John Wiley & Sons, 2011, vol. 27.

[28] W. Xu, M. Huang, C. Zhu, and A. Dammann, "Maximum likelihood TOA and OTDOA estimation with first arriving path detection for 3GPP LTE system," *Transactions on Emerging Telecommunications Technologies*, vol. 27, no. 3, pp. 339–356, Mar. 2016.

[29] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA): Requirements for support of radio resource management*, 3GPP TS 36.133, Rev. 14.13.0, Oct. 2019. [Online]. Available: https://www.3gpp.org/DynaReport/36133.htm

[30] D. Chase, "Code combining - a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," *IEEE Trans. on Communications*, vol. 33, no. 5, pp. 385–393, 1985.

[31] L.-A. Dufrene, M. Liu, M. Crussiere, J.-F. Helard, and J. Schwoerer, "Blind repetitions for cellular-IoT: Performance analysis of combination mechanisms," in *2016 23rd International Conference on Telecommunications (ICT)*, pp. 1–6.

[32] Y. Li and G. L. Stüber, *Orthogonal Frequency Division Multiplexing for Wireless Communications*. Springer, 2006.

[33] A. Ali and W. Hamouda, "Employing broadcast channel for frequency tracking in LTE-MTC systems," *IEEE Wireless Communications Letters*, vol. 5, no. 4, pp. 436–439, Jun. 2016.

[34] M. Speth, S. Fechtel, G. Fock, and H. Meyr, "Optimum receiver design for wireless broad-band systems using OFDM. I," *IEEE Trans. on Communications*, vol. 47, no. 11, pp. 1668–1677, Nov. 1999.

[35] C. Oberli, "ML-based tracking algorithms for MIMO-OFDM," *IEEE Trans. on Wireless Communications*, vol. 6, no. 7, pp. 2630–2639, Jul. 2007.

[36] M. Morelli and M. Moretti, "Fine carrier and sampling frequency synchronization in OFDM systems," *IEEE Trans. on Wireless Communications*, vol. 9, no. 4, pp. 1514–1524, Apr. 2010.

[37] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory.* Prentice-Hall PTR, 1993.

[38] D. Rife and R. Boorstyn, "Single-tone parameter estimation from discrete-time observations," *IEEE Trans. on Information Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.

[39] Q. Wang, C. Mehlführer, C. Mehlführer, and M. Rupp, "Carrier frequency synchronization in the downlink of 3GPP LTE," in *2010 21st Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 939–944.

[40] J. van de Beek, M. Sandell, and P. Borjesson, "ML estimation of time and frequency offset in OFDM systems," *IEEE Trans. on Signal Processing*, vol. 45, no. 7, pp. 1800–1805, Jul. 1997.

[41] K. Manolakis, A. Ibing, and V. Jungnickel, "Performance evaluation of a 3GPP LTE terminal receiver," in *2008 14th European Wireless Conference*, pp. 1–5.

[42] W. Xu and K. Manolakis, "Robust synchronization for 3GPP LTE system," in *2010 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–5.

[43] P. Frenger, P. Moberg, J. Malmodin, Y. Jading, and I. Godor, "Reducing energy consumption in LTE with cell DTX," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5.

[44] H. Kröll, M. Korb, B. Weber, S. Willi, and Q. Huang, "Maximum-likelihood detection for energy-efficient timing acquisition in NB-IoT," in *2017 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, pp. 1–5.

[45] M. Morelli, A. D'Andrea, and U. Mengali, "Feedback frequency synchronization for OFDM applications," *IEEE Communications Letters*, vol. 5, no. 1, pp. 28–30, Jan. 2001.

[46] K. Guo, W. Xu, and G. Zhou, "Differential carrier frequency offset and sampling frequency offset estimation for 3GPP LTE," in *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)*, pp. 1–5.

[47] MATLAB, *9.9.0 (R2020b)*. Natick, Massachusetts: The MathWorks Inc., 2020.

[48] MATLAB, "lteFadingChannel," Natick, Massachusetts, 2020. [Online]. Available: https://www.mathworks.com/help/lte/ref/ltefadingchannel.html

[49] J.-J. van de Beek, O. Edfors, M. Sandell, S. Wilson, and P. Borjesson, "On channel estimation in OFDM systems," in *1995 IEEE 45th Vehicular Technology Conference. Countdown to the Wireless Twenty-First Century*, vol. 2, pp. 815–819.

[50] M. Salomon, S. Lippuner, M. Korb, and Q. Huang, "Implementation and performance evaluation of cellular NB-IoT OTDOA positioning," in *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pp. 1365–1371.

[51] A. Hoglund, J. Bergman, X. Lin, O. Liberg, A. Ratilainen, H. S. Razaghi, T. Tirronen, and E. A. Yavuz, "Overview of 3GPP release 14 further enhanced MTC," *IEEE Communications Standards Magazine*, vol. 2, no. 2, pp. 84–89, Jul. 2018.

[52] A. Hoglund, X. Lin, O. Liberg, A. Behravan, E. A. Yavuz, M. Van Der Zee, Y. Sui, T. Tirronen, A. Ratilainen, and D. Eriksson, "Overview of 3GPP release 14 enhanced NB-IoT," *IEEE Network*, vol. 31, no. 6, pp. 16–22, Nov. 2017.

[53] 3GPP, *Evolved Universal Terrestrial Radio Access (E-UTRA): LTE Positioning Protocol (LPP)*, 3GPP TS 36.355, Rev. 14.7.0, Oct. 2018. [Online]. Available: https://www.3gpp.org/DynaReport/36355.htm

[54] X. Lin, J. Bergman, F. Gunnarsson, O. Liberg, S. M. Razavi, H. S. Razaghi, H. Rydn, and Y. Sui, "Positioning for the internet of things: A 3GPP perspective," *IEEE Communications Magazine*, vol. 55, no. 12, pp. 179–185, Sep. 2017.

[55] S. Hu, A. Berg, X. Li, and F. Rusek, "Improving the performance of OTDOA based positioning in NB-IoT systems," in *2017 IEEE Global Communications Conference (GLOBECOM)*, pp. 1–7.

[56] M. Korb, S. Willi, B. Weber, H. Kröll, A. Traber, S. Altorfer, D. Tschopp, J. Rogin, E. Dornbierer, M. Salomon, S. Lippuner, L. Wu, and Q. Huang, "A dual-mode NB-IoT and EC-GSM RF-SoC achieving -128-dBm extended-coverage and supporting OTDOA and A-GPS positioning," in *European Solid-State Circuits Conference (ESSCIRC)*, pp. 286–289.

[57] J. Lee, J. Han, C.-L. Lo, J. Lee, W. Kim, S. Kim, B. Kang, J. Han, S. Jung, T. Nomiyama, J. Lee, T. B. Cho, and I. Kang, "NB-IoT and GNSS all-in-one system-on-chip integrating RF transceiver, 23-dBm CMOS power amplifier, power management unit, and clock management system for low cost solution," *IEEE Journal of Solid-State Circuits*, vol. 55, no. 12, pp. 3400–3413, Aug. 2020.

[58] H. Guo, T. F. Chan, Y. T. Lai, K. C. Wan, L. Chen, and W. P. Wong, "30.3 a SAW-less NB-IoT RF transceiver with hybrid polar and on-chip switching PA supporting power class 3 multi-tone transmission," in *2020 IEEE International Solid-State Circuits Conference (ISSCC)*, pp. 464–466.

[59] ARM, "AMBA AXI and ACE Protocol Specification," Oct. 2011. [Online]. Available: https://developer.arm.com/documentation/ihi0022/e/

[60] ARM, "AMBA APB Protocol Specification," Aug. 2004. [Online]. Available: https://developer.arm.com/documentation/ihi0024/c/

[61] Quectel, *BG96 Hardware Design*, Jul. 2018.

[62] Nordic Semiconductor, *nRF9160 Product Specification*, v1.0.

[63] X. Lin, S. Rommer, S. Euler, E. A. Yavuz, and R. S. Karlsson, "5G from space: An overview of 3GPP non-terrestrial networks," *IEEE Communications Standards Magazine*, vol. 5, no. 4, pp. 147–153, Oct. 2021.

[64] O. Liberg, S. E. Löwenmark, S. Euler, B. Hofström, T. Khan, X. Lin, and J. Sedin, "Narrowband internet of things for non-terrestrial networks," *IEEE Communications Standards Magazine*, vol. 4, no. 4, pp. 49–55, Dec. 2020.

[65] T. A. Khan, X. Lin, S. E. Löwenmark, O. Liberg, S. Euler, J. Sedin, E. A. Yavuz, H. Shokri-Razaghi, and H.-L. Määttänen, "The Internet of things from space: Transforming LTE machine type communications for non-terrestrial networks," *IEEE Communications Standards Magazine*, vol. 6, no. 2, pp. 57–63, Jun. 2022.

[66] S. M. Kay, *Fundamentals of Statistical Signal Processing: Detection Theory*. Prentice-Hall PTR, 1998.