TECHNISCHE UNIVERSITÄT WIEN

institute of telecommunications

DISSERTATION

Multiplexing Services in 5G New Radio: Optimal Resource Allocation based on Mixed Numerology and Mini-slot Approach

This dissertation is submitted for the degree of "Doktor der technischen Wissenschaften"

Dipl.-Ing. Ljiljana Marijanović

Born Januar 24, 1990

Matrikelnummer:1528313

2020

Faculty of Electrical Engineering and Information Technology

TU Wien

Advisors:

Asst. Prof. Dipl.-Ing. Dr. techn. Stefan Schwarz

Institute of Telecommunications
Technische Universität Wien
Austria

Univ. Prof. Dipl.-Ing. Dr. techn. Markus Rupp

Institute of Telecommunications
Technische Universität Wien
Austria

Examiners:

Prof. Dr. Hüseyin Arslan

Electrical Engineering Department
University of South Florida
Tampa, Florida, USA

Assoc. Prof. Ing. Jiri Blumenstein, Ph.D.

Department of Radio Electronics
Brno University of Technology
Czech Republic

Abstract

In order to meet the diverse requirements imposed by a massive number of applications, the fifth generation (5G) New Radio (NR) Physical Layer (PHY) is designed to provide a highly flexible framework. This flexibility is made possible through a scalable numerology. The term numerology refers to the PHY waveform parametrization and allows the use of different subcarrier spacings, symbol and slot durations. In addition to an efficient support of various service requirements, employing a scalable numerology allows a better adjustment of the PHY waveform to different channel conditions, providing more robustness against channel variations.

Despite increased flexibility provided by multiplexing different numerologies, there is also a drawback of this concept, i.e., Internumerology Interference (INI) caused by non-orthogonal subcarriers between different numerologies. In this thesis, a closed-form expression of INI is derived for both, Cyclic Prefix (CP)-Orthogonal Frequency Division Multiplexing (OFDM) and Universal Filtered Multicarrier (UFMC) as a possible beyond 5G technology. In addition to INI, interference induced by doubly-selective channels as well as the channel estimation error are considered in this thesis.

Considering a multi-user mixed numerology scenario under different channel conditions, a novel optimization algorithm for joint numerology and resource allocation is proposed in the thesis. The proposed algorithm is first represented by the optimal Integer Linear Programming (ILP) solution. In order to reduce the computational complexity for large scale scenarios with many users and/or large transmission bandwidth, several less complex methods are proposed. The tradeoff between the performance and complexity is also discussed.

To multiplex different applications within the same band and thereby enhance the adaptability of the PHY, Third Generation Partnership Project (3GPP) proposes the mini-slot concept in addition to the mixed numerology concept. Employing these two concepts, this thesis proposes novel optimization algorithms for the joint numerology and resource allocation considering both, different channel conditions and user requirements. Emphasis is, in particular, placed on the achievement of Low Latency (LL) communications. Furthermore, the impact of UFMC on the performance is also discussed.

Kurzfassung

Um den mannigfaltigen Anforderungen, die von einer Vielzahl an Anwendungen gefordert werden, gerecht zu werden, wurde die 5G NR PHY so gestaltet, dass sie ein äußerst anpassungsfähiges Framework zur Verfügung stellt. Diese Anpassungsfähigkeit wird durch eine skalierbare Numerologie möglich gemacht. Der Begriff Numerologie bezieht sich auf die Parametrisierung der Wellenform der physikalischen Schicht und erlaubt die Verwendung unterschiedlicher Subträger-Abstände und Symbol- sowie Slot-Dauern. Neben der effizienten Unterstützung diverser Anforderungen unterschiedlicher Dienste erlaubt eine skalierbare Numerologie auch eine bessere Anpassung der Wellenformen an verschiedene Kanalbedingungen, und bietet so größere Robustheit gegen Kanalschwankungen.

Trotz der erhöhten Flexibilität, die durch das Multiplexen von verschiedenen Numerologien erreicht wird, kommt dieses Konzept auch mit einem Nachteil. Dieser besteht in INI, welche durch nicht-orthogonale Subträger zwischen den verschiedenen Numerologien verursacht wird. In dieser Arbeit wird ein geschlossener Ausdruck der INI hergeleitet, und zwar sowohl für OFDM als auch UFMC als potenzielle Nachfolger-Technologie in beyond 5G Funknetzen. Zusätzlich zur INI wird auch die Interferenz und der Kanalschätzfehler, die bei doppelt selektiven Kanälen auftreten, in dieser Arbeit betrachtet.

Unter der Annahme eines Mehr-Benutzer Szenarios mit gemischter Numerologie und unter Betrachtung unterschiedlicher Kanalbedingungen wird in dieser Arbeit ein neuartiger Optimierungs-Algorithmus zur gleichzeitigen Zuteilung von Numerologie und Ressourcen (im Zeit-Frequenz Raster) vorgestellt. Der vorgeschlagene Algorithmus wird zuerst durch die optimale ILP Lösung abgebildet. Um die Berechnungskomplexität für ausgedehnte Szenarien mit vielen Benutzern und/oder großer Bandbreite zu verringern, werden mehrere weniger komplexe Methoden vorgestellt. Auch der Kompromiss zwischen Performanz und Komplexität wird diskutiert.

Um unterschiedliche Anwendungen innerhalb des selben Bandes multiplexen zu können und somit die Anpassungsfähigkeit der PHY zu erweitern, schlägt das 3GPP ein Mini-Slot Konzept zusätzlich zur gemischten Numerologie vor. Unter Anwendung dieser zwei Konzepte werden in dieser Arbeit neuartige Optimierungs-Algorithmen für die gleichzeitige Zuweisung von Numerologie und Ressourcen unter der Berücksichtigung von einerseits unterschiedlichen Kanalzuständen und andererseits Benutzeranforderungen vorgestellt. Hierbei wird das Augenmerk insbesondere auf LL Kommunikation gelegt. Des Weiteren wird der Einfluss von UFMC auf die Performance diskutiert.

To my grandfather Milomir.

Acknowledgements

First and foremost, I would like to express my very great appreciation to my main advisor Asst. Prof. Dipl.-Ing. Dr.techn. Stefan Schwarz for all the knowledge he selflessly shared, for his vision and inspiring and stimulating conversations. His tremendous support and patient guidance throughout the years have been very much appreciated.

I would like to express my deep gratitude to my second advisor Univ.Prof. Dipl.-Ing. Dr.techn. Markus Rupp, who enabled me to become a member of his group, and without whom this thesis would never have come into existence. In addition, I am particularly grateful for many useful discussions and his invaluable insight over the years.

Special thanks to my examiners, Prof. Hüseyin Arslan and Assoc. Prof. Ing. Jiri Blumenstein, Ph.D. for their willingness to give their time so generously and for reviewing the thesis.

I would also like to extend my thanks to Christian Doppler Laboratory in cooperation with companies such as Nokia and A1, without whose financial help my thesis would not be possible. Special thanks to Gottfried Schnabl and Birger Haetty (Nokia, Ulm), who always reminded me of not only the theoretical significance, but the industrial application of my PhD thesis as well. Equally thanks to Waltraud Müllner from A1 for her support.

My sincere thanks go to my dear friends, one of the most treasured things I have acquired in the past five years of studying and working at TU Wien, except the thesis itself: Blanca, Jelena, Mariam, Samira, Martin, Bashar, Shrief, Osman and many others. My thanks also go to my beloved friends from Serbia and Bosnia, with whom I have shared all precious moments in my life and who are missed deeply.

Finally, I wish to thank my family for their support and encouragement that enabled me to become who I am today: my parents Goran and Gordana, as well as my sisters Biljana and Gorana, without whom I would never have fully experienced the joy of life. My last words of gratitude go to my loving husband Zeljko, who is and has been my strongest support and with whom I am on the path of fulfilling my lifelong dream - a family.

Contents

Chapter 1

Introduction

1.1 Motivation and Scope of the Thesis

Although the idea of mobile telephony emerged already in the early 20th century, the first mobile telephone service was proposed by AT&T in 1946. These were analog networks with manual switching. The first generation (1G) mobile telecommunications systems originated in the early 1980s, when the first widely used analog cellular system, called the Advanced Mobile Phone System (AMPS), was created. This generation was designed for voice transmission. It was characterized by bulky, power-consuming and short battery life equipment. The second generation (2G) was represented by a digital cellular network within Global System for Mobile Communications (GSM) standard in Europe and digital-AMPS standard in the USA. In addition to voice transmission, Short Message Service (SMS) was introduced for the first time. From then on, the evolution of the GSM system as the most accepted standard has brought a new range of services. Thus, a transfer of multimedia, as well as video streaming content, was additionally enabled in the third generation (3G) through Universal Mobile Telecommunications System (UMTS) standard. The transition to the fourth generation (4G) mobile communications and the introduction of Long Term Evolution (LTE) standard proposed by the Third Generation Partnership Project (3GPP), established new techniques, allowing higher speeds of transmission. In addition, it provided a large number of devices with a better coverage, as well as a reduced transmission latency.

In 2015, the era of the fifth generation (5G) mobile networks was entered within a global standard for New Radio (NR) access technology developed by 3GPP. A massive increase in the number of devices as well as traffic volume required the use of higher frequencies that had been very little or unused until then. Nevertheless, the spectrum below 6 GHz is still being predominantly used to support different types of applications. Hence, 5G operates on a wide range of frequencies, from several hundreds of MHz to the millimeter-wave bands. It is expected to cope with various requirements and services, and hence its air interface has to provide a highly flexible configuration. It encourages new functionalities in order to support human-centric and machine-centric applications [1, 2]. To cope with a large variety of demands of different applications, 3GPP identified three main use-case

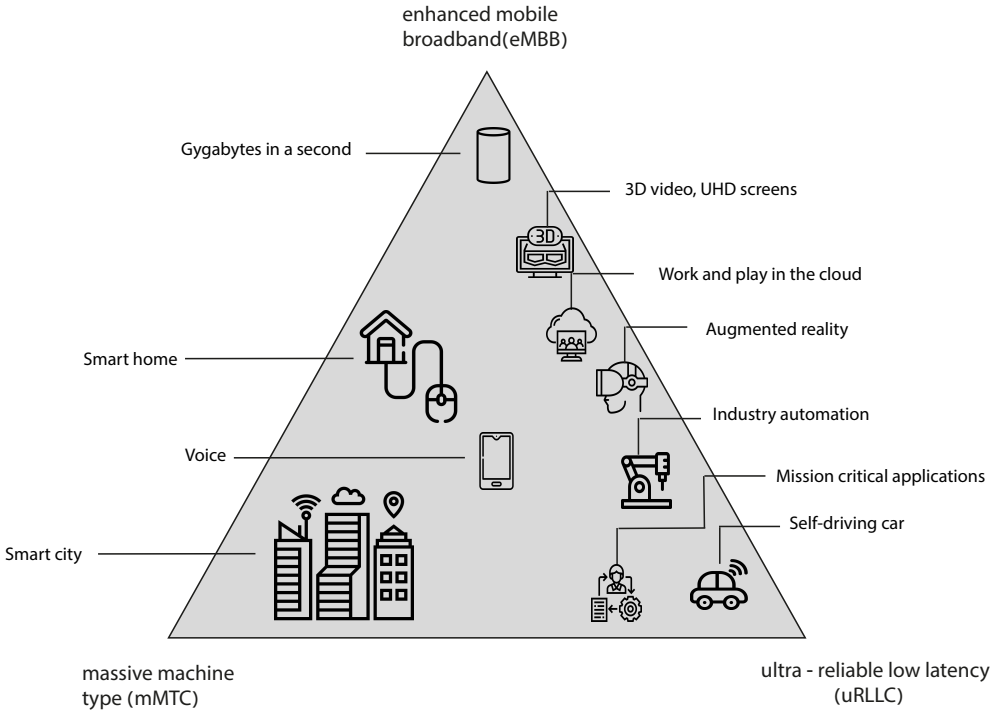


Figure 1.1: Use case families in 5G NR. ©ITU-R

families: Enhanced Mobile Broadband (eMBB), Massive Machine-Type Communications (mMTC) and Ultra-Reliable Low Latency Communications (uRLLC) [3, 4]. Thereby applications with the same or similar demands are united into one group as shown in Fig. 1.1. eMBB is a human-centric use case family that enables the access to multimedia content, services and data. It is envisioned to provide high data rates in high user density and low mobility scenarios on the one hand, and a wide area coverage in scenarios of lower user density but higher user mobility, on the other hand. The mMTC family addresses machine-centric use cases, supporting a massive number of connected devices. It is expected to transmit a small amount of traffic in non-critical delay scenarios, to have long battery life and to support low-cost devices. The uRLLC use case family caters to Low Latency (LL) sensitive applications such as remote medical surgery, autonomous driving, smart grids, industrial manufacturing, providing a highly reliable end-to-end link at the same time [5], [6].

To satisfy different requirements imposed by eMBB, mMTC and uRLLC traffic, 5G NR Physical Layer (PHY) has to provide a highly flexible framework [7]. The thesis is focused on multicarrier modulation component as the core component for the PHY design. This component implies both, employment of different NR waveforms and scalable numerology [8–10]. The first strategy considers the waveform contenders; among the many candidates in the 5G race, Cyclic Prefix (CP)-Orthogonal Frequency Division Multiplexing (OFDM) has been chosen by 3GPP as the most appropriate

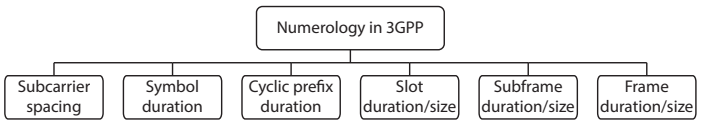


Figure 1.2: NR waveform parametrization.

waveform [11]. The suitability is reflected in the fact that OFDM has a good time localization which is crucial for LL communications and a low implementation complexity. In addition to OFDM, the thesis considers Universal Filtered Multicarrier (UFMC) as a potential beyond 5G technology. UFMC employs filtering on top of OFDM to improve spectral characteristics. The second strategy is a scalable numerology. Numerology refers to the waveform parametrization [2]. So far, there is only a single numerology specified in LTE downlink; however, as we approach a new generation this is enriched with more flexible structures. In Fig. 1.2 the set of flexible parameters proposed by 3GPP for 5G NR is shown. The crucial parameter is the subcarrier spacing. The other parameters are considered as dependent on it, simplifying the overall design. A joint configuration of the numerology and waveform is one of the biggest challenges not only for 5G, but also for the generations beyond. The importance of this adjustment is reflected in properly responding to the demands imposed by eMBB, mMTC and uRLLC [12–15].

To support multiple services on the same carrier, 3GPP proposes a mixed numerology for frequency multiplexing and a mini-slot approach that employs a single predefined numerology with shorter slot durations than a regular slot for that numerology [11]. In addition to an efficient support of various service requirements, employing a scalable numerology allows a better adjustment of the PHY waveform to different channel conditions, providing more robustness against channel variations. A scalable numerology enables the support of services with different requirements on the same carrier. One particular focus of 5G NR is on the achievement of LL communications [16–22]. Both mixed numerology and mini-slot methods can be used for LL transmissions. For example, a mini-slot approach, as the name implies, utilizes the resources within a time that is shorter than a regular slot duration, supporting LL and eMBB services at the same time. On the other hand, LL can also be accomplished employing wider subcarrier spacings, since in such a case slots have a shorter duration than an LTE slot.

Over the recent years, significant progress has been made in the context of resource management of flexible and mixed numerology transmissions. However, most of the research provides a heuristic approach to the problem, rarely producing a profound analysis and mathematical background of how the optimal numerology as well as the resources are chosen. In addition, they mostly consider only Quality of Service (QoS) requirements such as LL or high throughput imposed by the application, omitting the impact of pure PHY parameters, which is even more relevant in comparison to QoS demands. In the thesis, the optimal resource and numerology allocation for the mini-slot and mixed numerology approach is investigated, while taking into account an accurate PHY model as well as QoS demands.

1.2 Outline and Contributions

In this section, the main contribution of the thesis are emphasized throughout the following chapters.

Chapter 2: Preliminaries

In the race to 5G many waveform contenders were included. Considering the key performance indicators such as spectral efficiency, frequency and time localization, robustness to channel time and frequency selectivity, baseband complexity and many others, 3GPP selected CP-OFDM as well-suited waveform. However, for beyond 5G technologies there is also a possibility of filtering or windowing on the top of OFDM in order to improve the spectral efficiency. In the first part of this chapter, both CP-OFDM and UFMC are discussed. Throughout the step-wise analysis, an input-output relationship of both waveforms is provided. The second part of this chapter discusses the flexibility and scalability of these waveforms. In addition, the basic idea of the mini-slot and mixed numerology approaches for service multiplexing within one carrier is described.

Chapter 3: Flexible Numerology

An analysis of the PHY parameters that impact the numerology selection and formulates the Signal to Interference plus Noise Ratio (SINR) closed-form expression are provided in this chapter, employing both CP-OFDM and UFMC. Intercarrier Interference (ICI) and Intersymbol Interference (ISI) caused by channel variations over time and frequency are considered, as well as the inaccurate channel knowledge, due to imperfect channel estimation. Assuming a doubly-selective channel and accounting for the aforementioned imperfections, the optimal pilot pattern is selected for each multicarrier numerology. Exploiting the flexibility of 5G NR, the throughput performance is investigated and compared to LTE transmissions. The second part of this chapter discusses the optimal pilot pattern focusing on small data packet transmission which is typical for the mMTC use case. More precisely, the diamond-shaped pilot pattern already proposed by 3GPP in LTE is compared to a rectangular-shaped pilot pattern. The density and spacing of pilot symbols in both pilot patterns with respect to the constrained capacity is optimized and it is demonstrated that the rectangular pilot pattern can outperform the diamond-shaped pattern in case of small data packet transmissions.

The contributions of this chapter are published in the following papers:

[i] L. Marijanović, S. Schwarz, and M. Rupp, “Optimal numerology in OFDM systems based on imperfect channel knowledge”, in *87th Vehicular Technology Conference (VTC Spring)*, IEEE, 2018, pp. 1–5

[ii] L. Marijanović, S. Schwarz, and M. Rupp, "Pilot pattern optimization for small data packet transmission", in *26th Telecommunications Forum (TELFOR)*, IEEE, 2018, pp. 1–4

[iii] S. Pratschner, B. Tahir, L. Marijanović, M. Mussbah, K. Kirev, R. Nissel, S. Schwarz, and M. Rupp, "Versatile mobile communications simulation: The Vienna 5G Link Level Simulator", *EURASIP Journal on Wireless Communications and Networking*, 2018

[iv] S. Schwarz, B. R. Elbal, E. Zöchmann, L. Marijanović, and S. Pratschner, "Dependable wireless connectivity: Insights and methods for 5G and beyond", *e & i Elektrotechnik und Informationstechnik*, vol. 135, no. 7, pp. 449–455, 2018

Chapter 4: Investigation on Internumerology Interference

A scalable numerology offers increased flexibility on the one hand, but causes interference in scenarios of multi-user access with different subcarrier spacings on the other hand. In this chapter, a closed-form expression of Internumerology Interference (INI) is derived, employing both CP-OFDM and UFMC waveforms. The provided closed-form expressions include the subcarrier spacing, the filter length and the number of guard subcarriers between different numerologies. In addition, the analytical results are validated with simulation results. Introducing INI, the SINR expression has to be extended by this parameter. In order to follow the idea of the spectral mask proposed by 3GPP for the Out-of-Band (OOB) reduction, an upper bound on allowed INI is defined. By selecting the threshold, the guard band size between numerologies is determined.

The contributions of this chapter are published in the following papers:

[v] L. Marijanović, S. Schwarz, and M. Rupp, "Intercarrier interference of multiple access UFMC with flexible subcarrier spacings", in *25th European Signal Processing Conference (EUSIPCO), 2017*, IEEE, 2017, pp. 888–892

[vi] L. Marijanović, S. Schwarz, and M. Rupp, "Optimal resource allocation with flexible numerology", in *IEEE International Conference on Communication Systems (ICCS)*, IEEE, 2018, pp. 136–141

Chapter 5: Optimal Resource Allocation based on Mixed Numerology

In this chapter, a novel optimization method is proposed for joint resource and numerology allocation in a multi-user mixed numerology scenario. The optimal number of resources assigned to a certain user depends exclusively on the channel conditions that a user experiences. The optimization is given with respect to the frequency domain. The optimization can be formulated as an Integer Linear Programming (ILP). To reduce the computational complexity of the optimization, a linear relaxation of the problem is also proposed. The performance of the methods are investigated by numerical

simulations, demonstrating significant gains over an LTE-compliant scenario with fixed numerology and a heuristic approach.

The contributions of this chapter are published in the following papers:

[vii] L. Marijanović, S. Schwarz, and M. Rupp, "Optimal resource allocation with flexible numerology", in *IEEE International Conference on Communication Systems (ICCS)*, IEEE, 2018, pp. 136–141

[viii] L. Marijanović, S. Schwarz, and M. Rupp, "A novel optimization method for resource allocation based on mixed numerology", in *IEEE International Conference on Communications (ICC) 2019*, IEEE, 2019, pp. 1–6

Chapter 6: Optimal Resource Allocation based on Mixed Numerology and Mini-Slots for Low Latency Communications

In this chapter, the optimization investigated in the previous chapter is extended by accounting not only for the channel conditions, but also for the QoS demand, which is, in our case, latency. The focus is on the achievement of LL communications for a group of priority users, whereas the throughput of the other group of best effort users is maximized. In addition to the mixed numerology case, the constrained optimization for the mini-slot approach is formulated. Exploiting achievable rate performance as one of the fundamental metrics, these two approaches are compared under different channel conditions as well as different latency demands. The optimization problems for both mixed numerology and the mini-slot approach are initially given by an ILP solution. In order to reduce computational complexity for large-scale scenarios, the Dantzig-Wolfe decomposition method is also applied, showing that it is possible to achieve the optimal solution with significantly reduced complexity. The results are presented not only for CP-OFDM, but also for UFMC transmissions.

The contributions of this chapter are published in the following papers:

[ix] L. Marijanović, S. Schwarz, and M. Rupp, "Multi-user resource allocation for low latency communications based on mixed numerology", in *90th Vehicular Technology Conference (VTC2019-Fall)*, IEEE, 2019, pp. 1–7

[x] L. Marijanović, S. Schwarz, and M. Rupp, "Multiplexing services in 5G and beyond: Optimal resource allocation based on mixed numerology and mini-slots", vol. 8, 2020, pp. 209 537–209 555

Chapter 2

Preliminaries

2.1 Multicarrier Waveforms

In order to achieve a high degree of heterogeneity in terms of services and demands, many waveform contenders are postulated in the literature, such as: Filtered Bank Multicarrier (FBMC), Generalized Frequency Division Multiplexing (GFDM), UFMC, Weighted Overlap-and-Add (WOLA), filtered OFDM (f-OFDM) [32–36]. The waveforms are compared with respect to several key performance indicators including: time-frequency efficiency, frequency and time localization, robustness against time-frequency channel selectivity and robustness to carrier and time offset, Peak-to-Average-Power-Ratio (PAPR), Multiple Input Multiple Output (MIMO) compatibility, complexity and others [37–42]. Since CP-OFDM and UFMC waveforms are relevant for this work, in the following bulleted list a rough overview of their features is provided with respect to some of above-mentioned indicators.

- Time-frequency efficiency - The time-frequency efficiency is related to the achievement of high data rates and it is especially important for eMBB communications. In comparison to FBMC, the time-frequency efficiency of CP-OFDM and UFMC is reduced since they employ a CP and a Zero Padding (ZP), respectively [43].

- Frequency localization - CP-OFDM has a poor frequency localization and high OOB emissions due to its rectangular pulse shape. High frequency localization is especially beneficial when multiplexing different numerologies or waveforms simultaneously within the same carrier since it reduces interference between different numerologies, and therefore can operate with small guard bands. UFMC has a better frequency localization compared to CP-OFDM due to the filtering operation, but worse compared to FBMC [44].

- Time localization - A short symbol duration is needed to achieve uRLLC. Therefore, it is crucial to have a good waveform confinement in the time domain. Both CP-OFDM and UFMC are well- localized in the time domain and hence suitable for services that require a short time slot [45].

- Robustness against frequency-selective channels - Both CP-OFDM and UFMC are robust to frequency-selective channels since they employ a CP and a ZP, respectively. The robustness depends on many factors such as employed subcarrier spacing, signal bandwidth and cell deployment type.

- Robustness against time-selective channels - Rapid time variations of the wireless channel are expected in high mobility scenarios. Even though both OFDM and UFMC are relatively sensitive to the channel time variations, the robustness of these two waveforms can be improved by the selection of an appropriate subcarrier spacing [46].

- MIMO integration - MIMO transmission is vital for the communication systems since it can significantly improve spectral efficiency as well as coverage using beamforming techniques. Hence, it is important to utilize a waveform that is highly compatible with MIMO transmissions. Unlike FBMC that suffers from intrinsic interference, CP-OFDM and UFMC achieve orthogonality under perfect channel conditions. Therefore, these waveforms are better suited to MIMO transmission [47].

- PAPR - One of the main limitations of many modulation waveforms is a high PAPR. Unlike in LTE where Discrete Fourier Transform-Spread (DFTS)-OFDM is used for the uplink transmission due to low PAPR, in 5G NR there is an option to use either DFTS-OFDM or CP-OFDM. A low PAPR is required in order to meet a high energy efficiency demand. It can be achieved by applying conventional PAPR reduction techniques proposed for CP-OFDM, and some more sophisticated techniques proposed for UFMC [48–50].

- Waveform complexity - One of the most important features that influenced OFDM to be chosen as the waveform for 5G was certainly the transceiver baseband complexity. The complexity of CP-OFDM is quite low, whereas the complexity of UFMC is higher due to the filtering at the transmitter side and a larger Discrete Fourier Transform (DFT) size at the receiver side. However, there is also a proposed receiver design in the literature that achieves lower complexity [51, 52].

Low baseband complexity, very good compatibility with MIMO, the ability to combat time-frequency selective channels and good time localization are the features that make CP-OFDM a suitable waveform for the challenges in 5G. Other waveforms can be more or less suitable, depending on the requirements. Nevertheless, there is no single ideal waveform that performs always at the highest level. UFMC performs similarly to CP-OFDM, utilizing the spectrum more efficiently, but exhibiting higher complexity.

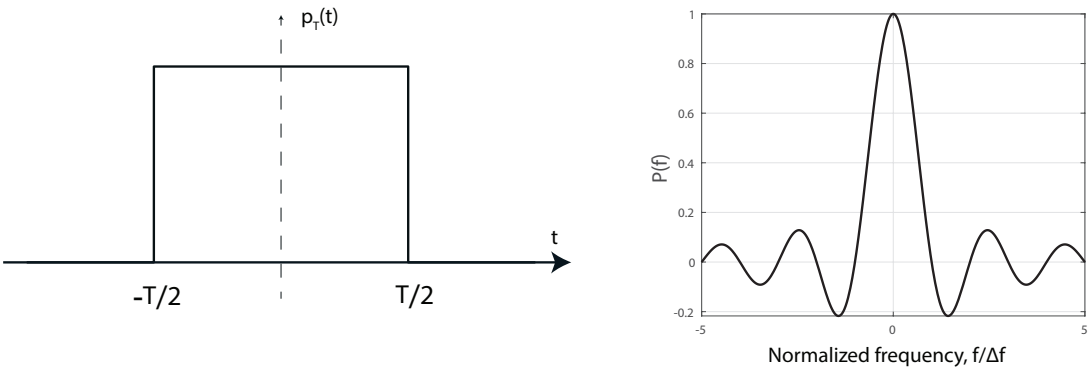


Figure 2.1: Rectangular pulse shape (left) and its frequency representation (right).

2.1.1 CP-OFDM

CP-OFDM is a modulation technique adopted in 4G LTE and, as previously mentioned, selected as the baseline transmission scheme of 5G NR. Multiple mutually orthogonal subcarriers transmit the information stream simultaneously over a range of frequencies, producing a high data rate system. CP-OFDM employs a rectangular pulse shape (prototype filter) $p_{\mathrm{T}}(t)$ in the time domain, resulting in a sinc shape $P(f)$ in the frequency domain, as shown in Fig. 2.1. Each of the subcarriers is represented by a shifted sinc spectrum at the frequency $f_k = f_i + k\Delta f$, with f_i being the initial frequency, $k \in \mathbb{N}_0$ being the subcarrier index and Δf is the subcarrier spacing. Multiple closely spaced orthogonal subcarriers decrease an extra overhead between adjacent subcarriers that is present in Frequency Division Multiplexing (FDM) and hence increase the spectral efficiency. The superposition of three subcarriers is illustrated in Fig. 2.2. By satisfying the condition of orthogonality, this modulation technique enables transmission without interference in ideal conditions.

The time domain OFDM signal is represented by the sum of the sine waves, each corresponding to a different frequency. The useful OFDM symbol duration is reciprocal to the subcarrier spacing Δf, i.e. $T_{\mathrm{u}} = 1/\Delta f$. In addition to the useful part, whose length corresponds to the number of subcarriers K, each OFDM symbol is extended by a CP of duration T_{CP}. The CP is represented by the last N_{CP} samples of the symbol that are copied in the front, preceding that symbol, as illustrated in Fig. 2.3. Thus, the total symbol length is $T = T_{\mathrm{CP}} + T_{\mathrm{u}}$. By employing the CP, a few main goals are achieved:

- The system is protected against the ISI induced by the frequency-selective multipath channel.
- Extra robustness is added to the receiver against the timing error.
- The frequency domain processing is simplified by enabling the cyclic convolution of the transmit data and the channel.

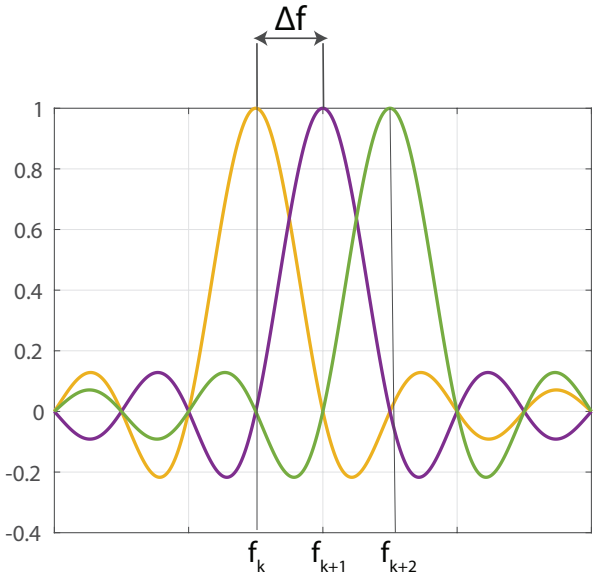


Figure 2.2: Spectra of three OFDM subcarriers.

The length of the CP is chosen to be 7 % of the symbol duration, with some exceptions in LTE and 5G NR that will be discussed later. The main disadvantage of this extra overhead is a reduced spectral efficiency and consequently a decreased data rate.

An equivalent baseband signal description of the OFDM signal transmission process is explained through the input-output relationship and this is supported by the transceiver block scheme shown in Fig. 2.4. Performing a serial-to-parallel transformation, K modulated symbols $\{X_k\}_{k=1}^{K}$ are mapped to the orthogonal subcarriers, and then a K point Inverse Discrete Fourier Transform (IDFT) operation is executed in order to obtain the time-domain samples $\{x_k\}_{k=1}^{K}$:

$$x_k = \mathfrak{F}^{-1}\{X_k\}, \quad k = 0, 1, ..., K. \tag{2.1}$$

The time-domain samples are then arranged in vector $\mathbf{x}$ via a parallel-to-serial transformation:

$$\mathbf{x} = [x_1, x_2, ..., x_K], \quad \mathbf{x} \in \mathbb{C}^{1\times K}. \tag{2.2}$$

The baseband transmit signal is extended by the CP:

$$\bar{\mathbf{x}} = [x_{K-N_{\mathrm{CP}}+1}, ..., x_K, x_1, x_2, ..., x_K], \quad \bar{\mathbf{x}} \in \mathbb{C}^{1\times(K+N_{\mathrm{CP}})}. \tag{2.3}$$

Assuming a time-invariant channel with channel impulse response $\mathbf{h}$, the extended baseband signal $\bar{\mathbf{x}}$ is then linearly convolved with the channel impulse response of length L_{CH} and impaired by the Additive White Gaussian Noise (AWGN), represented by vector $\mathbf{n} \sim \mathcal{CN}(0, \sigma_{\mathrm{n}}^{2,\mathrm{ofdm}})$, with noise variance $\sigma_{\mathrm{n}}^{2,\mathrm{ofdm}}$:

$$\bar{\mathbf{y}} = \bar{\mathbf{x}} * \mathbf{h} + \mathbf{n}, \tag{2.4}$$

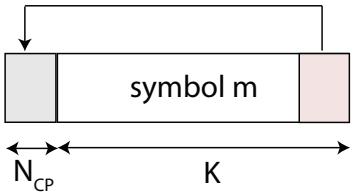


Figure 2.3: A structure of an OFDM symbol - length of a useful part of a symbol K and a CP length N_{CP} .

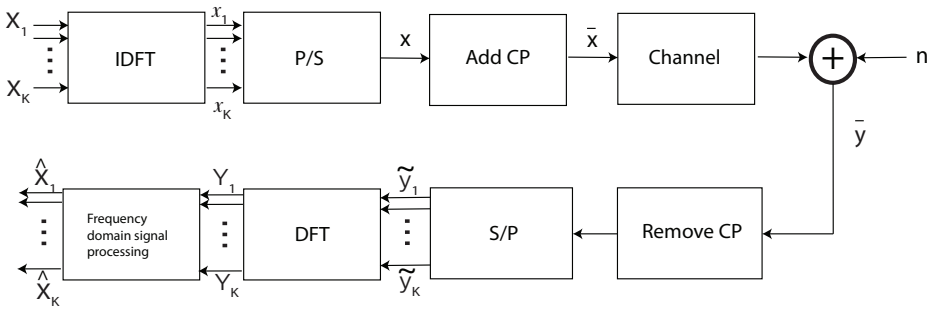


Figure 2.4: OFDM transceiver block diagram.

and $\bar{\mathbf{y}} \in \mathbb{C}^{1\times(K+N_{\mathrm{CP}}+L_{\mathrm{CH}}-1)}$ denoting the total received signal after the channel.

At the receiver the CP is removed and the process of the serial-to-parallel transformation is carried out:

$$\tilde{\mathbf{y}} = [\bar{y}_1, \bar{y}_2, ..., \bar{y}_K]^T, \quad \tilde{\mathbf{y}} \in \mathbb{C}^{K\times 1}. \tag{2.5}$$

Finally, the frequency domain representation of $\tilde{\mathbf{y}}$ is obtained by applying a K-DFT:

$$\mathbf{Y} = \mathfrak{F}\{\tilde{\mathbf{y}}\}, \quad \mathbf{Y} \in \mathbb{C}^{K\times 1}. \tag{2.6}$$

The presence of the CP enables the representation of the received signal in the time domain via a cyclic convolution between the transmit signal and the frequency response of the channel, resulting in the subcarrier-wise multiplication of these two parameters in the frequency domain. Hence, the output signal at subcarrier k, Y_k is given by:

$$Y_k = H_k X_k + N_k, \quad k \in \{1, 2, ..., K\}, \tag{2.7}$$

with N_k denoting the frequency domain noise.

As a final step, equalization is applied in order to estimate the transmitted signal $\hat{\mathbf{X}} = [\hat{X}_1, \hat{X}_2, ..., \hat{X}_K]$.

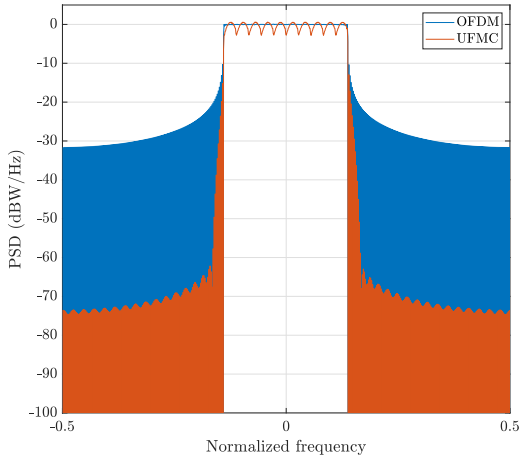


Figure 2.5: Power spectral density of OFDM and UFMC.

2.1.2 UFMC

UFMC is a modulation technique that was proposed for 5G NR. It is another in a series of techniques that applies filtering on top of the already existing structure employed by OFDM. UFMC offers flexibility in the sense that the assigned bandwidth can be divided into multiple subbands according to different user requirements or services, where each of the subbands is represented by a group of consecutive subcarriers. Applying a subband filtering, UFMC is well-suited for a fragmented spectrum and has a better frequency localization compared to OFDM, as shown in Fig. 2.5. This further implies reduced OOB emissions and increased robustness against ICI. Similarly to OFDM, the UFMC transceiver block scheme is shown in Fig. 2.6. Following this structure, the input-output relationship is derived. Vector $\mathbf{X}_b \in \mathbb{C}^{K\times 1}$ denotes the input symbols with active subband b, i.e.:

$$\mathbf{X}_b = [\mathbf{0}_{(b-1)K_b} X_{(b-1)K_b+1}, ..., X_{bK_b} \mathbf{0}_{K-bK_b-1}]^T, \tag{2.8}$$

where K_b is the total number of subcarriers per subband. Each of the subbands follows the conventional OFDM processing at the transmitter side. First, the frequency domain input signal is transformed to the time domain via a K-IDFT and a parallel-to-serial transformation is executed. The time domain signal is denoted as $\mathbf{x}_b \in \mathbb{C}^{1\times K}$.

Instead of appending a CP, a ZP of length N_{ZP} is inserted between consecutive symbols in UFMC in order to avoid ISI in case of a filter delay. The extended signal is denoted by $\mathbf{x}_b^{\mathrm{ZP}} \in \mathbb{C}^{1\times K+N_{\mathrm{ZP}}}$.There is also a possibility to avoid CP/ZP between symbols and therefore increase the spectral efficiency. However, in such cases it is needed to have a sophisticated algorithm for the ISI cancellation. Since a ZP does not have the same properties as a CP, a cyclic convolution is not preserved and a simple frequency multiplication of the transmit signal and channel does not remain.

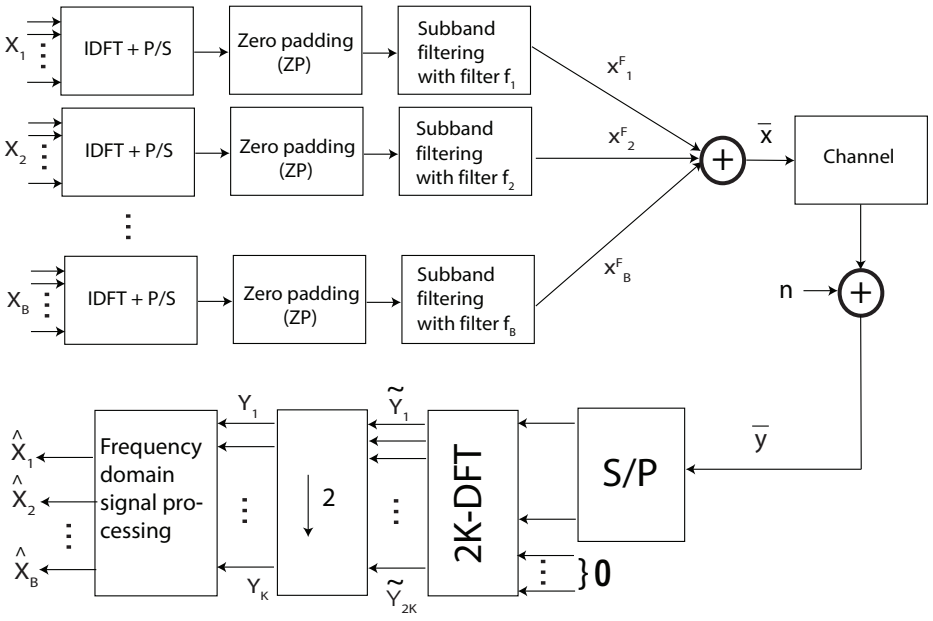


Figure 2.6: UFMC transceiver block scheme.

The main distinction of UFMC compared to CP-OFDM is a subband filtering with a Finite Impulse Response (FIR) filter:

$$\mathbf{f}_b^{\mathrm{p}} = \Big[f_b[0], f_b[1], ..., f_b[L_{\mathrm{F}}-1]\Big], \tag{2.9}$$

where L_{F} is the unique filter length applied for all subbands. The subband filter is a shifted version of the prototype filter on the subband center frequency f_{ce}:

$$f_b[i] = f_b^{\mathrm{p}}[i]\mathrm{exp}^{\frac{-j2\pi i f_{\mathrm{ce}}}{K}}. \tag{2.10}$$

Without loss of generality, it is assumed that the power of $\mathbf{f}_b$ is normalized to unity, i.e., $\sum_{i=0}^{L_{\mathrm{F}}} |f_b[i]|^2 = 1$. The signal after subband filtering is denoted by $\mathbf{x}_b^{\mathrm{F}} \in \mathbb{C}^{1\times(K+N_{\mathrm{ZP}}+L_{\mathrm{F}}-1)}$:

$$\mathbf{x}_b^{\mathrm{F}} = \mathbf{x}_b^{\mathrm{ZP}} * \mathbf{f}_b, \tag{2.11}$$

where $*$ denotes the convolution operator. There are several criteria for an optimal filter design in UFMC [53]. In this implementation a Hanning filter is employed since it shows very good side-lobe attenuation [54]. Additionally, a longer filter length means better spectral efficiency. However, there is a tradeoff between the spectral efficiency and intrinsic interference caused by a filter with greater length than the one of the ZP [55, 56]. As it is shown in Fig. 2.7, the filter length is the same as the ZP, in which case intrinsic interference does not exist.

The total transmit signal $\bar{\mathbf{x}} \in \mathbb{C}^{1\times(K+N_{\mathrm{ZP}}+L_{\mathrm{F}}-1)}$ is the sum of signals from individual subbands

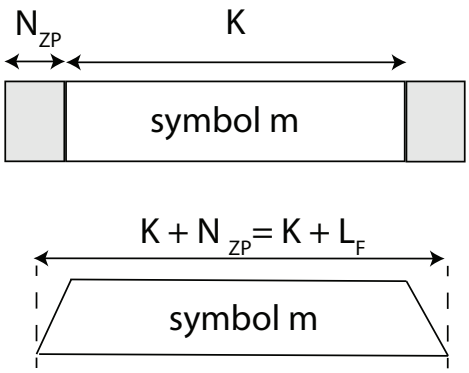


Figure 2.7: UFMC symbol structure where $L_{\mathrm{F}} = N_{\mathrm{ZP}}$.

(users):

$$\bar{\mathbf{x}} = \sum_{b=1}^{B} \mathbf{x}_b^{\mathrm{F}}, \tag{2.12}$$

where B is the total number of subbands. The signal is transmitted through the channel $\mathbf{h}$ of length L_{CH} and noise $\mathbf{n} \sim \mathcal{CN}(0, \sigma_{\mathrm{n}}^{2,\mathrm{ufmc}})$ is added:

$$\bar{\mathbf{y}} = \bar{\mathbf{x}} * \mathbf{h} + \mathbf{n}, \tag{2.13}$$

with slightly enhanced noise variance $\sigma_{\mathrm{n}}^{2,\mathrm{ufmc}}$ compared to the noise variance of CP-OFDM as defined in (3.14), and $\bar{\mathbf{y}} \in \mathbb{C}^{1\times(K+N_{\mathrm{ZP}}+L_{\mathrm{F}}+L_{\mathrm{CH}}-2)}$ is the received signal after the channel.

In addition to the ZP, there is another distinction that concerns the length of the DFT at the receiver side. There are two possibilities with respect to the length of the DFT: to use a K -DFT at the receiver side as described in [51] and keep the same complexity as in CP-OFDM, or to use a $2K$-DFT. The latter approach is mainly used in the literature for recovery of data symbols with a larger number of time domain samples compared to CP-OFDM. By employing this DFT size it is ensured that a cyclic convolution is preserved and therefore a simple subcarrier-wise multiplication can be used. A $2K$-DFT is performed at the receiver side by padding $K - L_{\mathrm{F}} - L_{\mathrm{CH}} + 2$ samples at the end of the received signal:

$$\tilde{\mathbf{Y}} = \mathfrak{F}\{\tilde{\mathbf{y}}\}, \tag{2.14}$$

where $\tilde{\mathbf{y}}$ is the extended received signal after a serial-to-parallel transformation and before the DFT operation, i.e., $\tilde{\mathbf{y}} = [\bar{\mathbf{y}}; \mathbf{0}_{(K-L_{\mathrm{F}}-L_{\mathrm{CH}}+2)\times 1}]$. In order to obtain useful data symbols out of $2K$ samples, the received signal $\tilde{\mathbf{Y}}$ is downsampled by a factor of two, so that only even samples are taken into account:

$$\mathbf{Y} = \tilde{\mathbf{Y}}^{\downarrow 2}. \tag{2.15}$$

Similarly to CP-OFDM, the equalization is performed and the estimated subband signal $\hat{\mathbf{X}}_b$ is extracted.

2.2 Scalable Numerology for New Radio

The need to support a large number of different requirements and services and provide a highly flexible PHY framework has led to the idea of using not just one numerology, as was the case in 4G LTE, but several. As already mentioned, the numerology refers to the waveform parametrization, i.e., a flexible employment of the parameters such as subcarrier spacing, symbol, slot and CP duration. In order to simplify the implementation, 3GPP in Release 15 proposed a set of numerologies that are scaled with respect to the baseline numerology which is 15 kHz subcarrier spacing [57]. Namely, the numerologies are scaled by the factor 2^{s-1}, where s is a positive integer (numerology index):

$$\Delta f_s = 2^{s-1}\Delta f_{s=1}, \tag{2.16}$$

and $\Delta f_{s=1}$ refers to the 15 kHz subcarrier spacing. The symbol duration is inversely proportional to the subcarrier spacing, i.e., $T_s = 1/\Delta f_s$. The relation of the numerologies with respect to the symbol duration is given by the following expression:

$$T_s = \frac{T_{s=1}}{2^{s-1}}, \tag{2.17}$$

with $T_{s=1}$ being the symbol duration that corresponds to 15 kHz subcarrier spacing, i.e, 66.67 μs. In the same way as the symbol duration, the CP duration is scaled:

$$T_{\mathrm{CP}_s} = \frac{T_{\mathrm{CP}_{s=1}}}{2^{s-1}}, \tag{2.18}$$

with $T_{\mathrm{CP}_{s=1}}$ denoting the CP duration of 15 kHz subcarrier spacing, i.e, 4.69 μs. Since larger subcarrier spacings are not suitable for many channel models with longer delay spreads due to their short CPs, 3GPP offers a possibility to use the extended CP for 60 kHz subcarrier spacing, which corresponds to 25 % of the total symbol duration.

The relevant numerology for this work is summarized in Table 2.1. The choice of scaling factor s

Table 2.1: NR flexible numerology

s	s = 1	s = 2	s = 3	s = 4
Subcarrier spacing [kHz]	15	30	60	120
Symbol duration [μs]	66.67	33.33	16.67	8.33
CP duration[μs]	4.69	2.34	1.17	0.59
Number of slots per subframe	1	2	4	8
Number of slots per frame	10	20	40	80
Frequency band [GHz]	0.45-6	0.45-6	0.45-6 24-52.6	24-52.6

depends on many aspects, such as carrier frequency, cell size, channel characteristics, latency and mobility. Large subcarrier spacings and thus shorter CP lengths are envisioned for small cell size, since that small cells experience lower delay spreads. Due to the propagation characteristics, small cell sizes and accordingly wider subcarrier spacings are well-suited at higher carrier frequencies. Smaller subcarrier spacings, on the other hand are relevant at lower carrier frequencies. In the thesis, the focus is on frequencies below 6 GHz; therefore, 15 kHz, 30 kHz and 60 kHz subcarrier spacings are employed. Going beyond the standardization framework, 120 kHz subcarrier spacing is also used.

Having different symbol and CP durations implies using different slot durations. According to 3GPP, a slot for each numerology is composed of 14 adjacent symbols of corresponding duration. Therefore, it can be accommodated more or fewer slots within a single subframe. Unlike in LTE where 2 slots of 0.5 ms compose one subframe for a 15 kHz subcarrier spacing, in 5G NR one slot corresponds to one subframe. The remaining values are listed in Table 2.1.

Depending on the requirements, it is possible to use either only one desired numerology or a combination of several. Mixing multiple numerologies led to the creation of a mixed numerology concept. For some applications that require a LL transmission, it is possible to use one desired numerology with a certain time span rather than multiple. That is how the mini-slot concept occurred. In the following two sections, an overview on mixed numerology and the mini-slot concept is provided.

2.2.1 Mixed Numerology

A mixed numerology concept implies the coexistence of multiple numerologies within one carrier in order to support numerous heterogeneous applications envisioned in 5G NR. The optimal choice of the numerology does not only depend on the service requirements (latency, reliability etc.) but also on the channel conditions imposed by the environment.

The unique choice of numerology cannot even be applied to a single use case family. For example, larger subcarrier spacings are required to achieve LL communications, since they provide shorter time slots. On the other hand, in a case when a channel delay spread is high, smaller subcarrier spacings ensure a high reliability as another important service requirement within the uRLLC use case family. The reason for this is that in such case the CP is longer, which prevents the system from ISI and reduces the probability of packet loss. The eMBB use case family seeks a high data rate, achievable at larger carrier frequencies, and exploiting larger subcarrier spacings. Large subcarrier spacings and a small number of subcarriers are preferred for the MTC use case family since it requires a short-data transmission and small packet sizes.

An example of mixed numerology is shown in Fig. 2.8. A general time-frequency resource allocation based on the minimum resource unit is shown here, represented by the Resource Element (RE). This

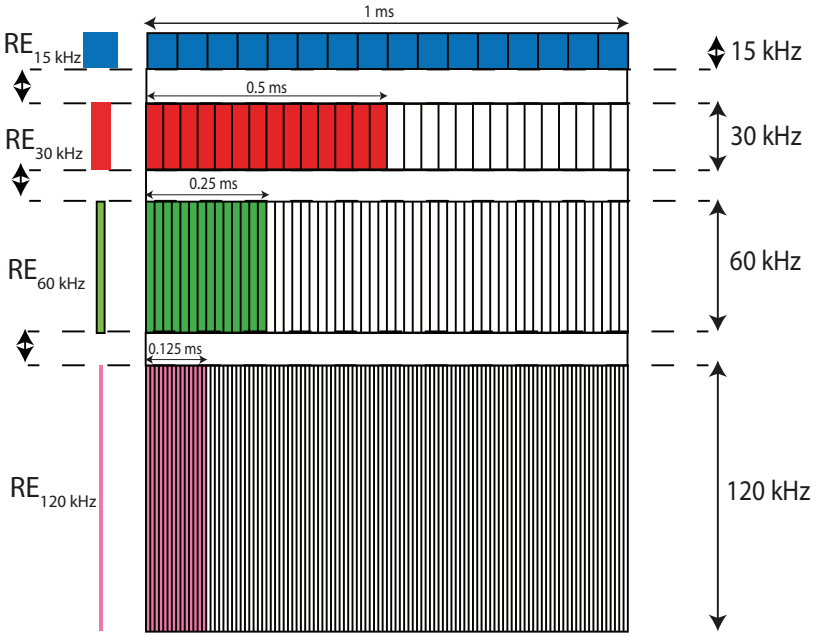


Figure 2.8: An example of a mixed numerology concept for all considered subcarrier spacings.

example is given for the subcarrier spacings from set $\mathcal{S}$, i.e., $\mathcal{S} = \{15, 30, 60, 120\}$ kHz. As can be noticed, a mixed numerology employs a regular slot duration in the time domain listed for each subcarrier spacing according to Table 2.1. The frequency domain representation is given with respect to one subcarrier of corresponding spacing; however, according to 3GPP the resource allocation is usually given with respect to the Resource Block (RB), i.e., 12 consecutive subcarriers.

Nevertheless, a mixed numerology has a major defect, which is the Internumerology Interference (INI) as a consequence of non-orthogonal subcarriers of different numerologies. In order to reduce this interference, an extra overhead in the sense of the Guard Band (GB) is employed. In Fig. 2.8 it is represented by white color. More details about this phenomenon will be discussed later.

2.2.2 Mini-slot

In addition to a mixed numerology approach for the frequency multiplexing of services, 3GPP proposed a mini-slot concept [11]. Unlike the mixed numerology that combines multiple numerologies together and implies a regular slot-based scheduling, a mini-slot concept employs a single predefined numerology with shorter slot durations than a regular slot for that numerology. It targets the similar use cases as the mixed numerology. LL communications as well as MTC are supported by using shorter time slots than a regular one, whereas longer slot durations are preferable for eMBB communications. Manifold benefits of a smaller slot durations are: providing a finer time domain multiplexing

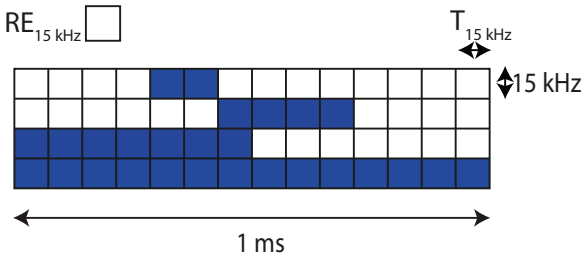


Figure 2.9: An example of a mini-slot concept for subcarrier spacing $\Delta f = 15\,\text{kHz}$.

granularity, transmissions in unlicensed spectrum as well as transmissions in the millimeter wave spectrum for scheduling users at shorter time scales [1].

According to 3GPP, a mini-slot can have a duration of 2, 4 or 7 consecutive symbols in the time domain. Assuming a single subcarrier of a certain spacing, it further implies 2, 4 or 7 REs, respectively, as shown in Fig. 2.9. It can be observed that the mini-slots can be positioned asynchronously with respect to the beginning of a regular slot. However, the boundaries of mini-slots have to be aligned with the symbol boundaries.

Chapter 3

Flexible Numerology

The term *flexibility* of numerologies for 5G NR has already been discussed in Chapter 2. The aim of this chapter is to show the gain that can be achieved by employing flexible numerologies compared to the previous generations. This chapter provides an insight into an optimal application of such numerologies under different channel conditions by investigating the numerologies independently. The changes that occur under different channel conditions are incorporated in the user post-equalization SINR closed-form expression and they are explained in detail in this chapter. The SINR expression accounts for ICI and ISI induced by a doubly-selective channel and a channel estimation error as a consequence of imperfect channel estimation. With the goal of maximizing the SINR, the optimal pilot positions for an individual numerology under specific channel conditions are selected. 3GPP has chosen the diamond pattern as applicable pilot pattern for 5G NR. However, under certain circumstances discussed in Section 3.3, it is shown that the diamond pilot pattern is not an optimal choice.

3.1 Post-equalization SINR Analysis

In Chapter 4 and Chapter 5 of this thesis, the numerology optimization problems that utilize the post-equalization SINR as main optimization criterion are proposed. As basis for these optimizations, a detailed analysis of the SINR closed-form expression is provided, for both, CP-OFDM and UFMC.

3.1.1 Post-equalization SINR for the CP-OFDM System

The input-output relationship of the CP-OFDM system is described in Section 2.1.1. It is given assuming only the presence of AWGN, without any other disturbances occurring due to the selective channel. However, in a real environment, imperfections must be taken into account due to the high speed of mobile terminals as well as multipath propagation. Considering all channel imperfections,

the input-output relationship of the system with respect to the frequency-time positions k, n can be written as:

$$Y_{k,n}^{\text{ofdm}} = H_{k,n} X_{k,n}^{\text{ofdm}} + W_{k,n}^{\text{ofdm}}, \tag{3.1}$$

where $X_{k,n}^{\text{ofdm}}$ is the transmitted data symbol, $H_{k,n}$ is the frequency response of the system and $Y_{k,n}^{\text{ofdm}}$ denotes the received data symbol. The additive disturbance is represented by $W_{k,n}^{\text{ofdm}}$ and it considers not only the noise, but also interference and channel estimation errors, as discussed below.

As previously mentioned, OFDM is an orthogonal system. However, sometimes the orthogonality cannot be preserved due to the sensitivity of the system to the Doppler shifts caused by the movement of the transmitter and the receiver. This results in mutually non-orthogonal subcarriers, i.e., ICI. Assuming Jakes' Doppler spectrum, the normalized ICI power[1] $\sigma_{\text{ICI}_{\Delta f}}^{2,\text{ofdm}}$ is expressed according to [58]:

$$\sigma_{\text{ICI}_{\Delta f}}^{2,\text{ofdm}} = 1 - \int_{-1}^{1} (1 - |x|) J_0(2\pi f_{\text{d}} T_{\Delta f}^{\text{ofdm}} x) dx, \tag{3.2}$$

with J_0 being the zeroth-order Bessel function of the first kind, representing the channel time autocorrelation function. The maximum Doppler shift is denoted by f_{d} and it is given by:

$$f_{\text{d}} = \frac{v_{\max} f_{\text{c}}}{c_0}, \tag{3.3}$$

where $v_{\max}$ is the maximum user velocity, f_{c} is the carrier frequency and c_0 is the speed of light. The ICI expression in (3.2) is given for the case of an infinite number of subcarriers, assuming evenly distributed power between subcarriers within the entire bandwidth. As discussed in Chapter 2, different subcarrier spacings imply different durations of the OFDM symbols, which directly impacts the ICI power as shown in Fig. 3.1. Obviously, a larger subcarrier spacing means a shorter duration of the symbol, and hence lower ICI for all user velocities.

Unlike ICI which is a consequence of the time-selective channel, ISI arises as a consequence of the frequency-selective channel. Namely, if the CP duration $T_{\text{CP},\Delta f}$ is not sufficient to cover the maximum delay spread of the channel $\xi_{\max}$, ISI occurs [59]. The ISI power depends on the Power Delay Profile (PDP) of the channel $\rho(\xi)$ and can be expressed as:

$$\sigma_{\text{ISI}_{\Delta f}}^{2,\text{ofdm}} = \frac{1}{T_{\Delta f}^{\text{ofdm}}} \int_{0}^{\xi_{\max} - T_{\text{CP},\Delta f}} \int_{T_{\text{CP},\Delta f}+t}^{\xi_{\max}} \rho(\xi) d\xi dt. \tag{3.4}$$

Again, the ISI power calculated in (3.4) is normalized with respect to the receive power per subcarrier. As shown in Fig. 3.2, ISI increases with a longer Root Mean Square (RMS) delay spread and larger subcarrier spacings, since $T_{\text{CP},\Delta f}$ is scaled inversely proportional to the subcarrier spacing.

As already emphasized in Chapter 2, in order to compensate for all the impairments caused by the channel and to recover the data at the receiver side, it is necessary to employ an equalization technique.

[1] The ICI power is normalized with respect to the receive power per subcarrier.

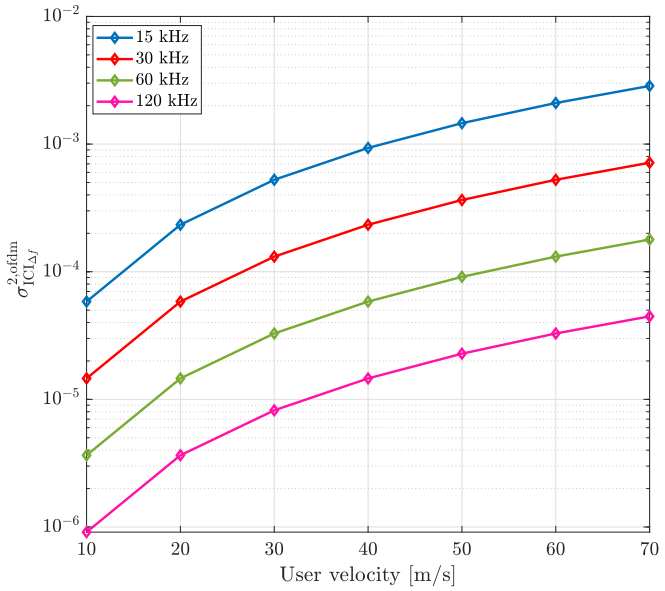


Figure 3.1: Intercarrier interference power vs. user velocity for considered subcarrier spacings.

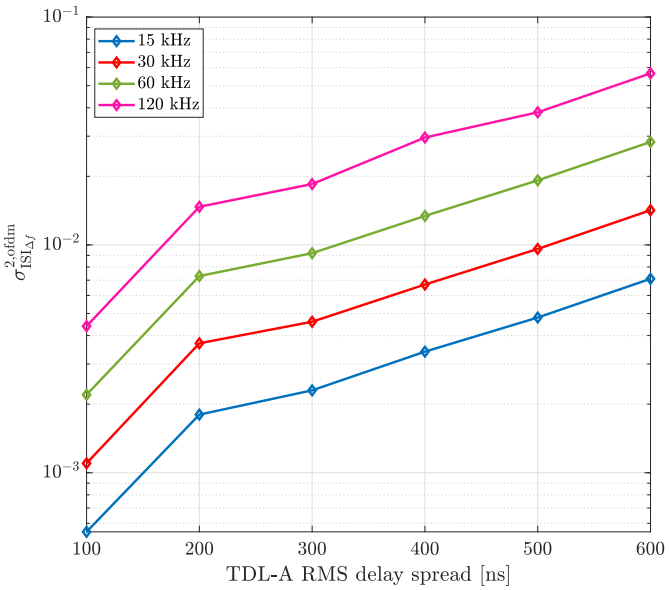


Figure 3.2: Intersymbol interference power vs. TDL-A RMS delay spread for considered subcarrier spacings.

Since the process of modulation and demodulation is performed on each subcarrier separately, the equalization process for CP-OFDM is greatly facilitated and can be performed by employing an one-tap equalizer.

In this thesis, the channel estimation is based on pilot symbols. An equal power distribution between the data and the pilot symbols is considered:

$$\begin{aligned} \sigma_d^2 &= \mathbb{E}[X_{k,n}]^2 = 1 \\ \sigma_p^2 &= \mathbb{E}[X_{k_p,n_p}]^2 = 1, \end{aligned} \tag{3.5}$$

with $k_p, n_p \in \mathcal{P}$ denoting the frequency-time pilot position from set $\mathcal{P}$ of all pilot positions. First, the channel is estimated at the pilot positions, performing a Least Square (LS) estimation:

$$\hat{H}_{k_p,n_p} = \frac{Y_{k_p,n_p}}{X_{k_p,n_p}}. \tag{3.6}$$

Given these estimates, a two dimensional linear interpolation is applied to estimate the channel at the data positions $\hat{H}_{k,n}$:

$$\hat{H}_{k,n} = \sum_{k_p,n_p \in \mathcal{P}} w_{k,n}^{k_p,n_p} \hat{H}_{k_p,n_p}, \tag{3.7}$$

with inter/extrapolation weights $w_{k,n}^{k_p,n_p}$ [23], [60], [61].

Due to the imperfect channel estimation and therefore an inaccurate channel knowledge, it is necessary to take into account the channel estimation error. The Mean Squared Error (MSE) caused by LS channel estimation can be expressed as [60]:

$$\sigma_{\mathrm{e}_{\Delta f}}^{2,\mathrm{ofdm}} = c_{\mathrm{e},\Delta f}\left(\sigma_{\mathrm{n}}^{2,\mathrm{ofdm}} + \sigma_{\mathrm{ICI}_{\Delta f}}^{2,\mathrm{ofdm}} + \sigma_{\mathrm{ISI}_{\Delta f}}^{2,\mathrm{ofdm}}\right) + d_{\mathrm{e},\Delta f}. \tag{3.8}$$

In addition to ICI and ISI, the error depends on scalars $c_{\mathrm{e},\Delta f}$ and $d_{\mathrm{e},\Delta f}$. Scalar $c_{\mathrm{e},\Delta f}$ is determined by the weighting factor of (3.7):

$$c_{\mathrm{e},\Delta f} = \frac{1}{N_{\mathrm{d}}} \sum_{\{k,n\} \in \mathfrak{D}} \sum_{\{k_p,n_p\} \in \mathcal{P}} \left(w_{k,n}^{k_p,n_p}\right)^2. \tag{3.9}$$

This scalar is the average value over all data symbols N_{d} from data set $\mathfrak{D}$, depending only on the pilot pattern. Scalar $d_{\mathrm{e},\Delta f}$ refers to the interpolation error and it depends not only on the pilot positions, but also on the second order statistics of the channel as captured by autocorrelation function $\mathbf{R}_{\Delta k,\Delta n}$:

$$\begin{aligned} d_{\mathrm{e},\Delta f} = \frac{1}{N_{\mathrm{d}}} \sum_{\{k,n\} \in \mathfrak{D}} \Bigg(1 - 2 \sum_{\{k_p,n_p\} \in \mathcal{P}} w_{k,n}^{k_p,n_p} \Re\{\mathbf{R}_{\Delta k,\Delta n}\} \\ + \sum_{\{k_p,n_p\} \in \mathcal{P}} \sum_{\{k_{p'},n_{p'}\} \in \mathcal{P}} w_{k,n}^{k_p,n_p} w_{k,n}^{k_{p'},n_{p'}} \Re\{\mathbf{R}_{\Delta k_{pp'},\Delta n_{pp'}}\}\Bigg). \end{aligned} \tag{3.10}$$

The indices Δk and $\Delta k_{pp'}$ refer to the absolute distance between the data and the pilots in the frequency domain, respectively. The distances in the time domain, Δn and $\Delta n_{pp'}$, are defined in the same way. By considering the channel under the Wide Sense Stationary Uncorrelated Scattering (WSSUS) approximation, the autocorrelation function is defined as the Kronecker product ($\otimes$) of temporal $\mathbf{R}_{\Delta n}$ and frequency correlation $\mathbf{R}_{\Delta k}$:

$$\mathbf{R}_{\Delta k,\Delta n} = \mathbf{R}_{\Delta n} \otimes \mathbf{R}_{\Delta k}. \tag{3.11}$$

This arises from the fact that the statistics of such a channel do not change over time and frequency, and hence its transfer function is a 2D stationary process [62]. The Kronecker structure of the channel autocorrelation function corresponds to independent temporal and frequency correlation. The temporal correlation is modeled assuming the Jakes' spectrum and frequency correlation is determined by the PDP of the channel. Intuitively, stronger frequency and time correlation imply smaller interpolation error.

Finally, by taking into account all the aforementioned parameters, the post-equalization SINR under imperfect channel knowledge is given by:

$$\begin{aligned} \bar{\gamma}_{\Delta f}^{\mathrm{ofdm}} &= \frac{\mathbb{E}[X_{k,n}]^2}{\mathbb{E}[\hat{X}_{k,n} - X_{k,n}]^2} \\ &= \frac{\sigma_{\mathrm{d}}^2}{\sigma_{\mathrm{n}}^{2,\mathrm{ofdm}} + \sigma_{\mathrm{ICI}_{\Delta f}}^{2,\mathrm{ofdm}} + \sigma_{\mathrm{ISI}_{\Delta f}}^{2,\mathrm{ofdm}} + \sigma_{\mathrm{e}_{\Delta f}}^{2,\mathrm{ofdm}}}. \end{aligned} \tag{3.12}$$

3.1.2 Post-equalization SINR for the UFMC System

Unlike CP-OFDM that discards a CP at the receiver side in order to remove the potential ISI introduced by the channel, UFMC does not discard a ZP, but utilizes $K + N_{\mathrm{ZP}} + L_{\mathrm{F}} - 1$ received samples. Due to the extended signal length, a $2K$-DFTs is performed in order to recover the transmit data. The input-output relationship of a single user UFMC transmission is given by:

$$\tilde{Y}_{k,n}^{\mathrm{ufmc}} = \tilde{H}_{k,n} \sum_{s} \tilde{X}_{k,n}^{b} \tilde{F}_{k,n}^{b} + \tilde{W}_{k,n}^{\mathrm{ufmc}}, \tag{3.13}$$

where $\tilde{H}_{k,n}$ and $\tilde{F}_{k,n}^{b}$ are $2K$-DFTs of the channel response and the time domain subband filter, respectively. The $2K$-DFT of the data symbols is denoted by $\tilde{X}_{k,n}^{b}$. The last term $\tilde{W}_{k,n}^{\mathrm{ufmc}}$ comprises ISI and ICI as the channel properties, AWGN $\sim \mathcal{CN}(0, \sigma_{\mathrm{n}}^{2,\mathrm{ufmc}})$ with slightly increased noise variance compared to the time domain white noise and channel estimation error $\sigma_{\mathrm{e}_{\Delta f}}^{2,\mathrm{ufmc}}$. It should be recalled that only even samples from (3.13) are relevant for data detection.

There are many different implementation schemes for the UFMC technique proposed in the literature. In order to avoid unnecessary overhead, many authors employ neither a CP or a ZP. Instead, some

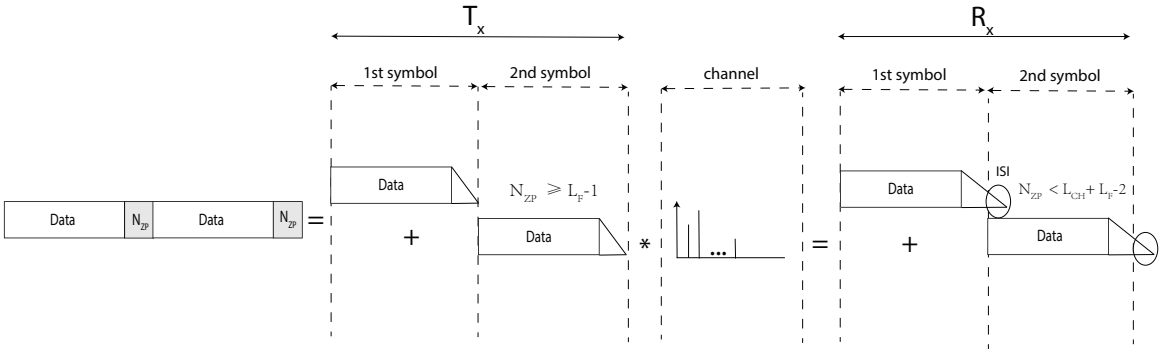


Figure 3.3: Illustration of UFMC transmission considering two symbols.

authors utilize a filter ramp-up and ramp-down as a soft protection against highly frequency-selective channels, reducing the ISI power [63]. Theoretically, there are two conditions that need to be fulfilled in order to ensure ISI-free transmission, as illustrated in Fig. 3.3:

- In the absence of the channel, the ZP length must cover the length of the filter, i.e. $N_{\mathrm{ZP}} \geq L_{\mathrm{F}} - 1$. Otherwise, ISI between two consecutive symbols is caused by the filter tail.
- In order to avoid ISI in case of a time dispersive channel, the combined length of the channel and filter length has to be shorter than the ZP length, i.e. $N_{\mathrm{ZP}} \geq L_{\mathrm{CH}} + L_{\mathrm{F}} - 2$.

Similarly to ISI, ICI can occur due to two main reasons:

- Inappropriate filter length breaks the orthogonality between the consecutive subbands b, especially violating the edge subcarriers.
- Highly time-selective channels cause non-orthogonality between the subcarriers.

If ICI and ISI exist, they are modelled as $\sigma^{2,\mathrm{ufmc}}_{\mathrm{ICI}}$ and $\sigma^{2,\mathrm{ufmc}}_{\mathrm{ISI}}$ according to [64].

The noise variance in UFMC, $\sigma^{2,\mathrm{ufmc}}_{\mathrm{n}}$, differs from the noise variance in CP-OFDM. It is slightly enhanced due to the increased number of the time domain samples as a consequence of the filtering:

$$\sigma^{2,\mathrm{ufmc}}_{\mathrm{n}} = \frac{K + L_{\mathrm{F}} - 1}{K} \sigma^{2,\mathrm{ofdm}}_{\mathrm{n}}. \tag{3.14}$$

Furthermore, there is a negligible noise correlation between different subcarriers.

In order to take the distortions introduced by the filter into account, a pre-equalization method can be performed at the transmitter side. Thereby, the channel estimation of OFDM can be fully applied to UFMC, without additional modifications. As can be seen from the previous description of the UFMC transceiver, this is not considered in this work because possible filter distortions are compensated by a sufficiently long ZP. Furthermore, the pilot placement in OFDM is completely reusable in UFMC [65].

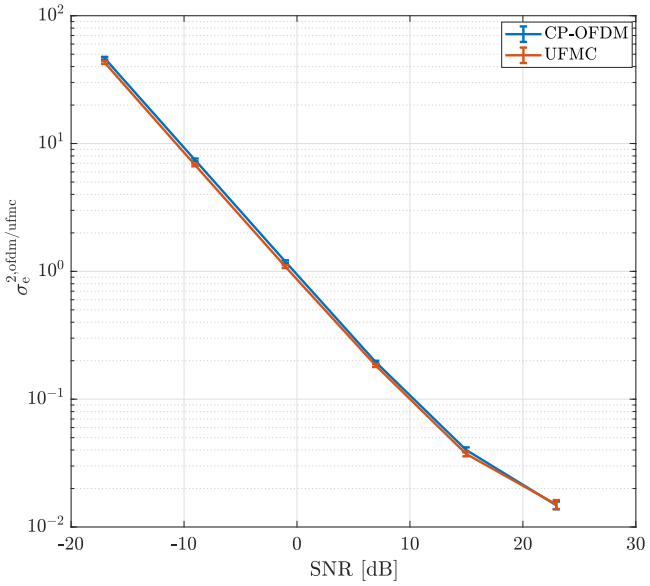


Figure 3.4: Comparison of the channel estimation error of CP-OFDM and UFMC.

In Fig. 3.4, the channel estimation errors of CP-OFDM and UFMC versus different Signal to Noise Ratio (SNR) values are compared. For these simulations the user velocity of 60 m/s and Tapped Delay Line (TDL)-A channel model with RMS delay spread of ξ = 350 ns are considered. It can be observed that the channel estimation errors of these two waveforms quite match along all SNR values.

The final SINR expression given in (3.12) can be translated to UFMC, considering the appropriate parameters for this technique.

$$\bar{\gamma}_{\Delta f}^{\mathrm{ufmc}} = \frac{\sigma_{\mathrm{d}}^2}{\sigma_{\mathrm{n}}^{2,\mathrm{ufmc}} + \sigma_{\mathrm{ICI}_{\Delta f}}^{2,\mathrm{ufmc}} + \sigma_{\mathrm{ISI}_{\Delta f}}^{2,\mathrm{ufmc}} + \sigma_{\mathrm{e}_{\Delta f}}^{2,\mathrm{ufmc}}}.$$

3.2 Performance Investigation of Flexible OFDM Numerology

The ICI and ISI performances of different OFDM multicarrier numerologies obtained according to the analytical expressions (3.2) and (3.4), are shown and explained in Section 3.1.1. In the following part, these theoretical analyses are approved by simulations obtained by employing the *Vienna 5G Link Level Simulator* [25].

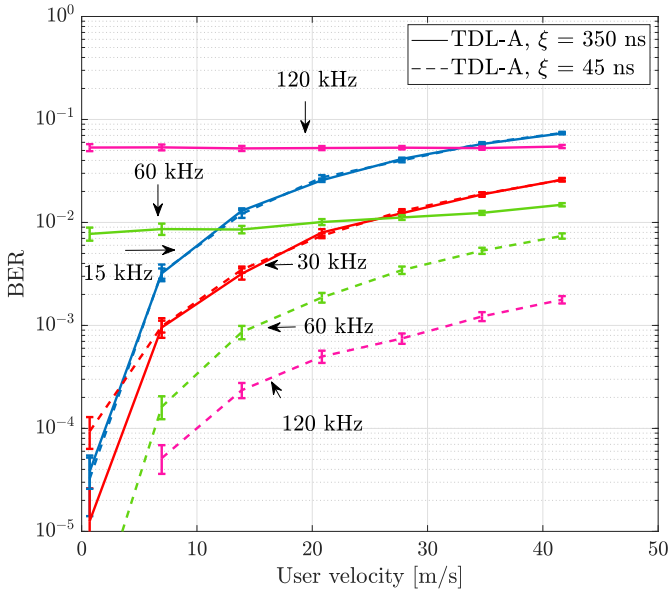


Figure 3.5: BER vs. user velocity for considered subcarrier spacings and two different channel RMS delay spreads.

3.2.1 Performance Investigation under Perfect Channel Knowledge

The aim of this subsection is to investigate the sensitivity of different numerologies with respect to delay and Doppler spread of the channel under perfect channel knowledge. In Fig. 3.5 different OFDM numerologies are compared with respect to the uncoded bit-error ratio (BER) considering the TDL-A channel model with two different RMS delay spreads, $\xi = 45$ ns and $\xi = 350$ ns, for a set of velocities shown in the x-axis [66]. The carrier frequency in this case is 5.9 GHz.

As investigated by the analysis, a better behaviour of larger subcarrier spacings, with respect to the Doppler shift, compared to smaller subcarrier spacings is confirmed by simulations only for $\xi = 45$ ns. The reason for this is that the larger subcarrier spacings experience less ICI. In this case, all subcarrier spacings experience rather small ISI. The largest subcarrier spacing of 120 kHz performs best in this case, since it experiences the lowest ICI. In general, it can be observed that the BER increases with the user velocity due to a growing ICI.

Considering the impact of the RMS delay spread, it is visible from Fig. 3.5 that smaller subcarrier spacings perform better, since they experience less ISI. A significant performance degradation is observed for 60 kHz and 120 kHz subcarrier spacings for $\xi = 350$ ns due to the insufficient CP length compared to the channel delay spread. The performances for 15 kHz subcarrier spacing for both RMS delay spread values match quite well due to the longer CP. The same holds true for 30 kHz

subcarrier spacing. However, this is only true when considering a perfect channel knowledge. With imperfect channel estimation, their performances degrade with an increase in the RMS delay spread, due to a growing channel estimation error.

3.2.2 Performance Investigation under Imperfect Channel Knowledge

Considering a real scenario without perfect channel knowledge, the aim of this section is to investigate the channel estimation error and throughput performance of different OFDM multicarrier numerologies, exploiting the flexibility offered in 5G. As emphasized, the channel estimation is based on pilot symbols. In the LTE standard, only a fixed pilot pattern with predefined pilot positions is proposed. However, taking into consideration doubly selective channels with imperfect channel knowledge, it is unreasonable to use the same fixed spacing between the pilots for all channel conditions [67]. Furthermore, the fixed pilot positions are not convenient for all numerologies, since numerologies have differently arranged time-frequency resources on the one hand and they are not equally adaptable to the same channel conditions on the other hand. Hence, the goal here is to find an adaptive pilot-symbol spacing in time D_{t} and frequency D_{f} for each numerology under different channel conditions, considering a diamond pilot pattern.

Optimization Problem Formulation

In order to find optimal pilot-symbol spacings and thereby maximize the throughput performance, the upper bound of the constrained capacity metric is employed as an objective function and it is defined as:

$$\bar{C}_{D_{\mathrm{f}},D_{\mathrm{t}},\Delta f} = B_{D_{\mathrm{f}},D_{\mathrm{t}}} \log_2(1 + \bar{\gamma}_{D_{\mathrm{f}},D_{\mathrm{t}},\Delta f}). \tag{3.15}$$

This objective function accounts for the average post-equalization SINR $\bar{\gamma}_{D_{\mathrm{f}},D_{\mathrm{t}},\Delta f}$ given by (3.12)[2] and the transmission bandwidth $B_{D_{\mathrm{f}},D_{\mathrm{t}}}$. The average post-equalization SINR metric considers the channel conditions and the estimation error due to the imperfect channel estimation. Intuitively, the SINR improves by decreasing the distance between the pilots since the estimation is more precise. However, accounting only for the average SINR means a smaller channel estimation error and a very small useful bandwidth. Therefore, the bandwidth utilization has to also be a part of the optimization. A bandwidth is represented as the ratio between the number of the transmitted data and the total number of resources:

$$B_{D_{\mathrm{f}},D_{\mathrm{t}}} = \frac{N_{\mathrm{d}}}{N_{\mathrm{d}} + N_{\mathrm{p}}}, \tag{3.16}$$

with N_{p} denoting the total number of pilot positions in a set $\mathcal{P}$. Both parameters N_{d} and N_{p} depend on parameters D_{f} and D_{t}, i.e., the smaller the pilot spacing in time and/or frequency, the less useful

[2]Notice that indices D_{f}, D_{t} are omitted in the average SINR expression defined by (3.12) for brevity.

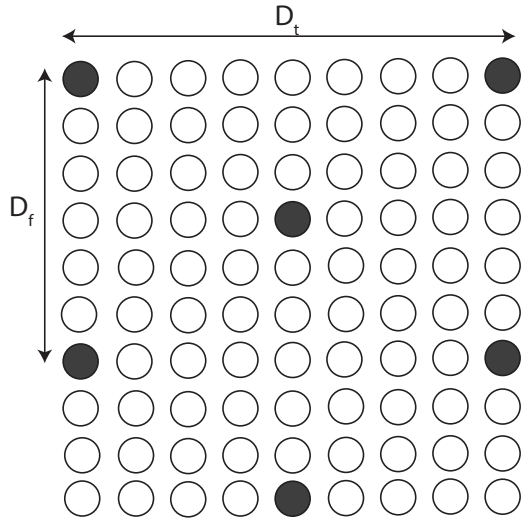


Figure 3.6: Pilot-symbol pattern with D_f = 6 and D_t = 8.

data symbols are available. Finally, the optimization problem can be written as follows:

$$\begin{aligned} \underset{D_\mathrm{f},D_\mathrm{t},\Delta f}{\text{maximize}} \quad & \bar{C}_{D_\mathrm{f},D_\mathrm{t},\Delta f} \\ \text{subject to} \quad & \Delta f \in 15 \cdot 2^{s-1}\text{kHz}, s \in \mathcal{S} \\ & D_\mathrm{f} \in \{2, ..., c_1\}, D_\mathrm{f} \in \mathbb{N}, c_1 \in \mathbb{N}, \\ & D_\mathrm{t} \in \{2, ..., c_2\}, D_\mathrm{t} \in \mathbb{N}, c_2 \in \mathbb{N}, \end{aligned} \tag{3.17}$$

where c_1 is a constant that depends on the total number of subcarriers and c_2 is a constant that depends on the total number of symbols within 1 ms. To solve this combinatorial optimization problem, the objective function is maximized by exhaustive search over all feasible combinations of the variables D_f and D_t for given subcarrier spacing Δf, defined by a subband s.

Analytical and Simulation Results

The MSE of the channel estimator as well as the expressions for the average weighting factor $c_{\mathrm{e},\Delta f}$ and the interpolation error $d_{\mathrm{e},\Delta f}$ are described in Section 3.1.1. The following outlines the dependence of these factors on different pilot spacings in frequency D_f and time D_t obtained by applying (3.8).

Similarly to LTE, the diamond-shaped pilot pattern has been selected for 5G NR [68]. In Fig. 3.6 an example of the diamond pilot pattern is shown. The data positions are represented by white dots and the pilots are represented by black dots. In order to account for different channel variations and to

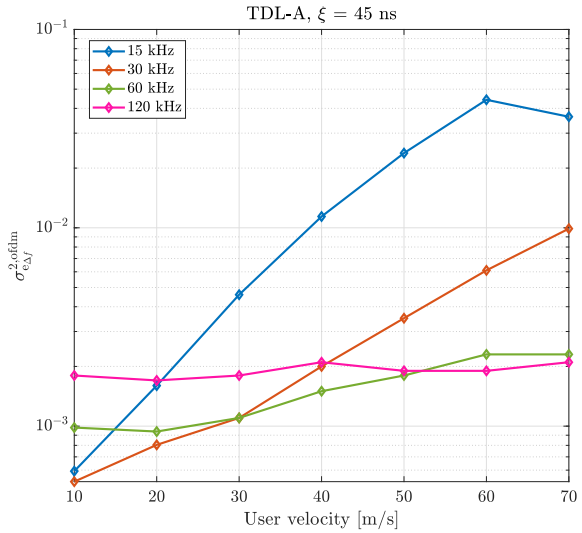


Figure 3.7: Channel estimation error with different numerologies - TDL-A with $\xi = 45$ ns.

decrease the MSE of the channel estimator, the pilot positions from the grid shown in Fig. 3.6 can be adapted.

- For a low Doppler spread and thus a large coherence time, the pilot spacing in the time domain should be increased. Conversely, for a low delay spread channel the coherence bandwidth is larger, and hence the pilot spacing in the frequency domain has to be increased.

- Under the same channel conditions, the pilot positions differ between different numerologies. Since there is a different time-frequency arrangement of resources of corresponding numerologies, different time-frequency density of pilots can be noticed. Although there is a proportionality between the waveform parameters among different numerologies, the proportionality between the optimal pilot positions and therefore the total number of pilots are not preserved.

Fig. 3.7 shows the dependence of the channel estimation error on user velocity applying the optimal pilot pattern for the particular Doppler shifts, assuming TDL-A channel model with $\xi = 45$ ns [69]. The behaviour of the curves is a consequence of the time-frequency correlation tradeoff. Namely, the impact of temporal correlation is noticeable with lower subcarrier spacings, that is, the temporal correlation is stronger and the interpolation error is smaller with smaller user velocity. However, when observing the curves with 60 kHz and 120 kHz subcarrier spacings, the frequency correlation is more dominant at all velocities, which results in performance saturation of these cases. Additionally, a dominance of ISI over ICI and interpolation error is already visible with $\xi = 45$ ns for larger subcarrier spacings. Due to a very large ISI power as well as the interpolation error in the case of

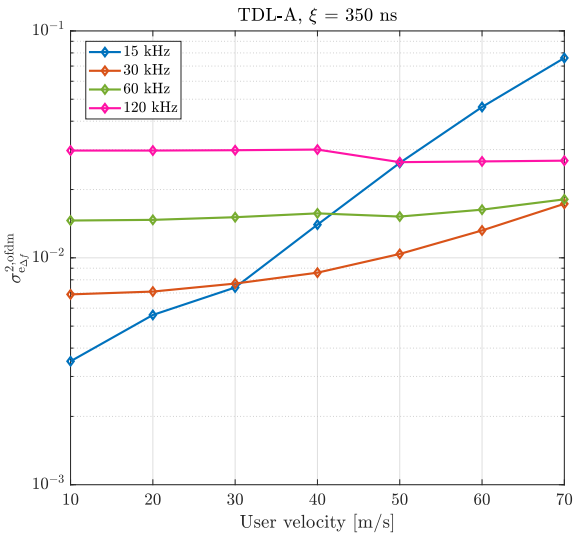


Figure 3.8: Channel estimation error with different numerologies - TDL-A with $\xi = 350$ ns.

TDL-A channel with RMS delay spread $\xi = 350$ ns [69], the larger subcarrier spacings show even larger gap in comparison to the lower subcarrier spacings as shown in Fig. 3.8.

The objective function defined in (3.17) is utilized in order to predict the throughput performances. The throughput is defined as the number of successfully decoded data within a frame per frame unit. For a constant SNR of 30 dB, a certain velocity and a particular subcarrier spacing, the optimal pilot distances in the frequency and time domain are selected. A bandwidth of 1.44 MHz is used for each numerology, but the number of subcarriers and the symbols changes accordingly.

The throughput performances with respect to the user velocity for TDL-A channel model with RMS delay spreads of $\xi = 45$ ns and $\xi = 350$ ns are shown in Fig. 3.9 and Fig. 3.10, respectively. With a larger subcarrier spacing for TDL-A channel of $\xi = 45$ ns the larger throughput is achieved, since the larger subcarrier spacing is more robust to ICI and also shows good performance in terms of interpolation error. Additionally, the LTE case with a fixed pilot spacings is compared to a baseline numerology of 15 kHz with the optimal pilot pattern for each velocity, keeping the same CP length in both cases. In these circumstances, it is visible that the optimized case outperforms the LTE performance.

On the other hand, from Fig. 3.10 it can be observed that a larger subcarrier spacing does not always mean a better throughput performance due to the more pronounced differences between frequency and time correlation with larger subcarrier spacings. Due to very high interpolation error and ISI larger subcarrier spacings show very poor performance. For this scenario it is noticeable that the optimal subcarrier spacing varies for different velocities. Also, it can be observed that the channel delay spread has more impact on the performance than the Doppler shift.

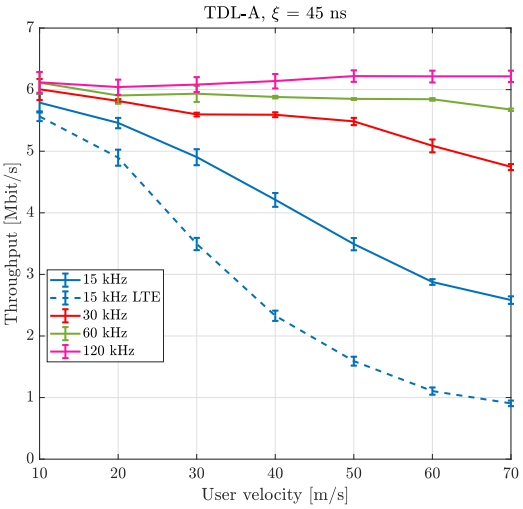


Figure 3.9: Throughput with different numerologies - TDL-A with $\xi = 45$ ns.

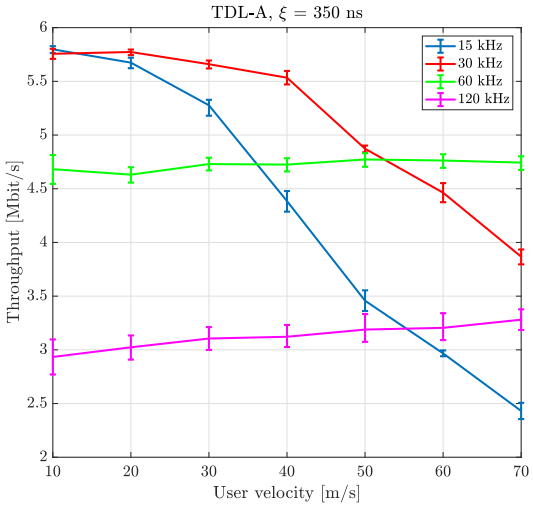
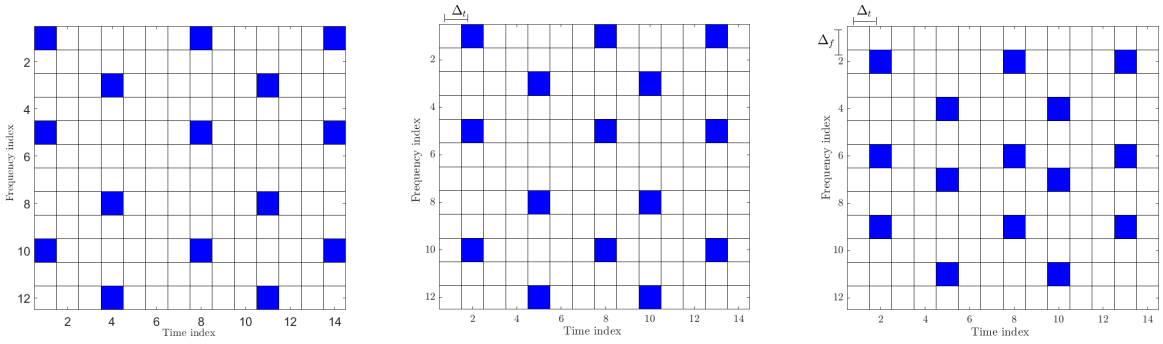


Figure 3.10: Throughput with different numerologies - TDL-A with $\xi = 350$ ns.

3.3 Pilot Pattern Optimization for MTC

In the literature, a diamond pilot pattern or its variants, such as a rotated diamond or a cell pattern are presented as optimal, under the assumption of unbounded transmission bandwidth [68]. In this case, only the interpolation effect between pilots has to be considered and extrapolation effects at the borders are negligible. However, in reality, with a finite bandwidth the optimal pilot pattern also depends on the extrapolation error that is involved in the channel estimation error. The extrapolation error depends on the number of data resources that lie outside the grid of pilot symbols. Generally, placing the pilots closer to the border results in a smaller extrapolation error. In case of a large transmit bandwidth, the impact of the extrapolation error is negligible; however, this does not apply to a small transmit bandwidth. Small transmission bandwidths are typical for the MTC use case family. The focus of this section is to demonstrate the optimal placement of diamond-shaped and rectangular-shaped pilots with possibly unequal pilot spacing within different circumstances.

In order to show that the optimality of the diamond pilot pattern is not always preserved, it is compared to a rectangular-shaped pilot pattern under 15 kHz subcarrier spacing. In contrast to the common assumption of equally spaced pilots in the time and frequency domains, which is a reasonable assumption for the unbounded transmission bandwidth and duration, unequally spaced pilots are considered here, achieving a smaller extrapolation error. In order to minimize the channel



(a) **(b)** **(c)**

Figure 3.11: Pilot shifts in the time and frequency domain for a constant bandwidth: $(\mathbf{a})$ no shifts, $(\mathbf{b})$ $\Delta_\mathrm{t} = 1$, $\Delta_\mathrm{f} = 0$, $(\mathbf{c})$ $\Delta_\mathrm{t} = 1$, $\Delta_\mathrm{f} = 1$.

estimation error, so far the optimal pilot spacings in time D_t and frequency domain D_f are investigated. Additionally, in order to evaluate the extrapolation error, it is necessary to investigate the optimal pilot placement with respect to the borders of the transmit block with a constant number of pilots and data in time and frequency. First the pilots are placed at the borders of the transmission block, and then they are shifted symmetrically towards the middle and an extrapolation error is calculated at each step. The pilots are shifted either particularly in time, Δ_t, or frequency domain, Δ_f, or in both domains simultaneously, as it is shown in Fig. 3.11. Fig. 3.11a represents the situation when pilots

are placed at the border of the transmit block, while Fig. 3.11b and Fig. 3.11c are cases with only time shift and shifts in both domains, respectively. This figure shows the diamond pilot pattern for one RB for 15 kHz subcarrier spacing. The same setup holds for the rectangular pilot pattern.

The optimal time-frequency shifts vary depending on the channel conditions and bandwidth size. To optimize the spacing of the pilot symbols and hence the total number of pilots in the frequency domain, $N_{\mathrm{p_f}}$, and in the time domain, $N_{\mathrm{p_t}}$, as well as the time-frequency shifts, the upper bound of the constrained capacity as an objective function is employed [60]:

$$\bar{C}_{N_{\mathrm{p_f}},N_{\mathrm{p_t}},\Delta_{\mathrm{f}},\Delta_{\mathrm{t}},\Delta f} = B_{N_{\mathrm{p_f}},N_{\mathrm{p_t}}} \log_2(1+\bar{\gamma}_{N_{\mathrm{p_f}},N_{\mathrm{p_t}},\Delta_{\mathrm{f}},\Delta_{\mathrm{t}},\Delta f}), \tag{3.18}$$

where $\bar{\gamma}_{N_{\mathrm{p_f}},N_{\mathrm{p_t}},\Delta_{\mathrm{f}},\Delta_{\mathrm{t}},\Delta f}$ is the post-equalization SINR that additionally depends on the time-frequency shifts. The goal is to maximize the upper bound of the constrained capacity and find the optimal number of pilots in the frequency and time domain. It can be defined as:

$$\begin{aligned}
&\underset{N_{\mathrm{p_f}},N_{\mathrm{p_t}},\Delta_{\mathrm{f}},\Delta_{\mathrm{t}}}{\text{maximize}} \quad \bar{C}_{N_{\mathrm{p_f}},N_{\mathrm{p_t}},\Delta_{\mathrm{f}},\Delta_{\mathrm{t}}} \\
&\text{subject to} \quad \Delta f = 15\text{kHz}, \\
&\qquad D_{\mathrm{f}} \in \{2,...,c_1\}, D_{\mathrm{f}} \in \mathbb{N}, c_1 \in \mathbb{N}, \\
&\qquad D_{\mathrm{t}} \in \{2,...,c_2\}, D_{\mathrm{t}} \in \mathbb{N}, c_2 \in \mathbb{N}, \\
&\qquad \Delta_{\mathrm{f}} \in \{0,...,c_3\}, \Delta_{\mathrm{f}} \in \mathbb{N}, c_3 \in \mathbb{N}, \\
&\qquad \Delta_{\mathrm{t}} \in \{0,...,c_4\}, \Delta_{\mathrm{t}} \in \mathbb{N}, c_4 \in \mathbb{N},
\end{aligned} \tag{3.19}$$

where c_3 and c_4 are the constants that depend on the block size and the number of pilots in time and frequency. In performed simulations, these variables take values from the range $\{0,1,2,3\}$.

In Fig. 3.12 the behaviour of the constrained capacity versus the number of pilots in the frequency domain is shown. Two different velocities of 5 km/h for TDL-A channel model with ξ = 45 ns and 150 km/h for TDL-A channel model with ξ = 250 ns are considered. One resource block for 15 kHz numerology (180 kHz bandwidth) is observed. The SNR is considered to be 20 dB. The results are based on the optimal number of pilots in time and the optimal time-frequency shifts at the same time. In both cases it can be observed that the rectangular pilot pattern outperforms the diamond pilot pattern due to smaller extrapolation error.

In order to clarify this, Fig. 3.13 is provided. This specific case considers TDL-A channel model with ξ = 45 ns and user velocity of 5 km/h. Moreover, it considers two pilot symbols in the frequency domain and the optimal number of pilots in the time domain, as well as, the optimal time-frequency shifts for both, the diamond-shaped pilot pattern and the rectangular-shaped pilot pattern. The rectangular-shaped pilot pattern shown in Fig. 3.13a achieves the maximum capacity for two pilots in the time domain. For the same number of pilots in the time domain the diamond-shaped pilot pattern

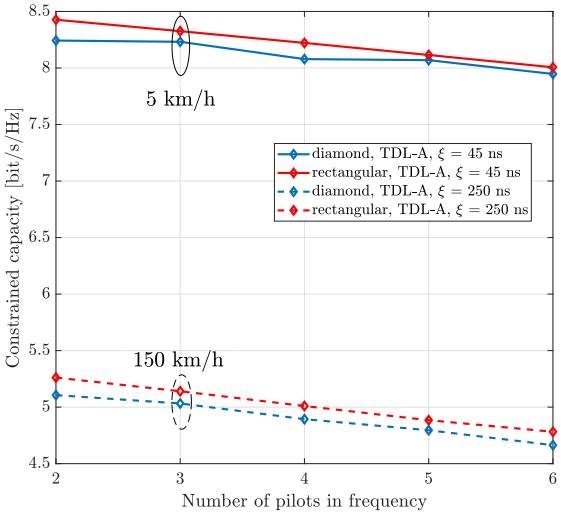
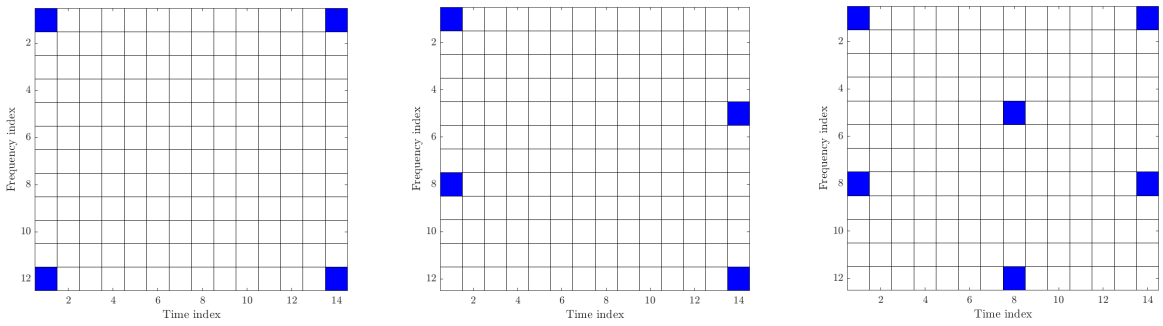


Figure 3.12: Constrained capacity vs. number of pilots in the frequency domain.

shown in Fig. 3.13b suffers from larger channel estimation error compared to the rectangular-shaped pilot pattern. The reason for that are larger border effects, mainly represented through $c_{\mathrm{e},\Delta f}$ defined in (3.9). The diamond-shaped pilot pattern shown in Fig. 3.13c achieves the maximum capacity for three pilots in the time domain due to the smallest channel estimation error. Additionally, due to the increased number of pilots and hence a smaller useful bandwidth, this case results in a smaller capacity compared to the rectangular pilot placement shown in Fig. 3.13a.



(a) **(b)** **(c)**

Figure 3.13: Illustration of different pilot placements and border effects for $N_{\mathrm{f}} = 2$ and $(\mathbf{a})$ $N_{\mathrm{t}} = 2$ with the rectangular-shaped pilot pattern, $(\mathbf{b})$ $N_{\mathrm{t}} = 2$ with the diamond-shaped pilot pattern and $(\mathbf{c})$ $N_{\mathrm{t}} = 3$ with the diamond-shaped pilot pattern.

Similarly, in Fig. 3.14 it is presented how the constrained capacity changes with respect to the number

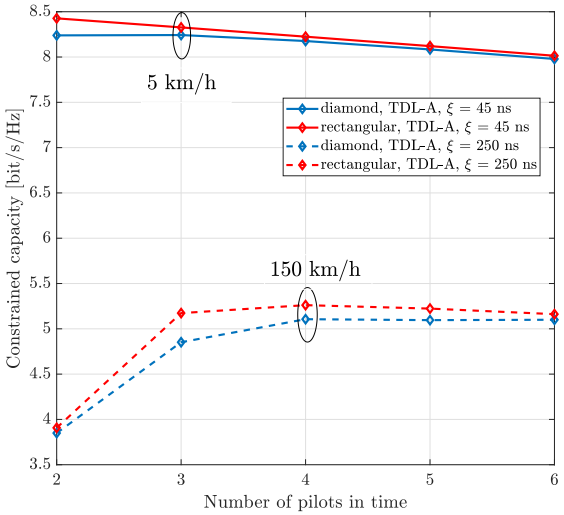


Figure 3.14: Constrained capacity vs. number of pilots in the time domain.

of pilots in the time domain. Similarly to the previous example, the rectangular pilot pattern shows better performances compared to the diamond pilot pattern. It can be observed that more pilot symbols in the time domain are needed to achieve the maximum capacity for both pilot patterns.

Fig. 3.15 shows how the constrained capacity changes with respect to different velocities for the optimal configuration of parameters. Again, a similar conclusion can be drawn as in previous examples. In case when TDL-A channel model with RMS delay spread of 45 ns is employed, the channel estimation error is dominant over both, ISI and ICI. For a larger RMS delay spread of 350 ns, the channel estimation error is still larger than a cumulative impact of ICI and ISI. However, the gap between different pilot patterns is decreased, resulting in their similar behaviour.

In order to verify the validity of the constrained capacity expression defined in (3.18), accounting also for the extrapolation error, the throughput performance is simulated. Here, a larger bandwidth of 72 subcarriers is used for both the analysis and the simulations. For a certain number of pilots in the frequency and the time domain the maximum throughput is found. The results are simulated over a finite set of modulation and coding schemes defined in LTE/5G and the best one is selected. From Fig. 3.16 it is visible that the simulated throughput performance has the same trend as the constrained capacity. The throughput and capacity performance are normalized with respect to their maximum achieved values marked in the same figure. It can be observed that the curves are almost constant over a large regime of pilots in the time and frequency grid. In this area a larger overhead is compensated by an SINR gain or, more precisely, by a small interpolation error due to a large number of pilots. However, once the overhead gets too large, the capacity performance starts deteriorating.

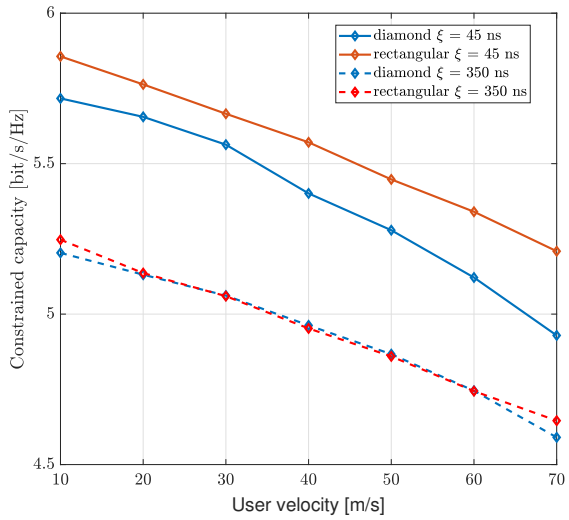


Figure 3.15: Constrained capacity vs. user velocity.

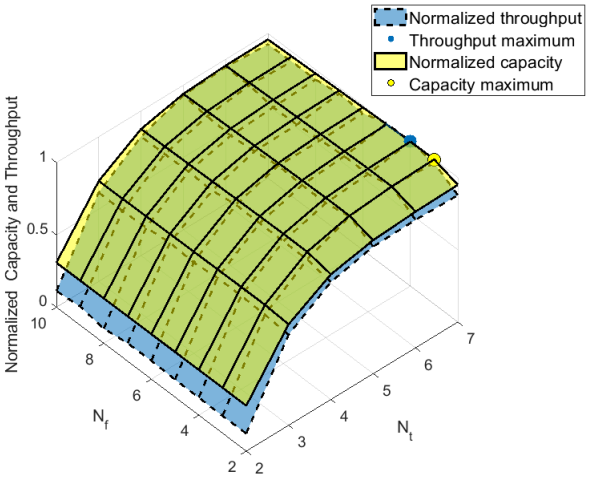


Figure 3.16: Comparison of normalized constrained capacity and normalized throughput.

Chapter 4

Investigation of Internumerology Interference

The main advantage of mixed numerology is that it provides increased flexibility and adaptability of the PHY; however, it also has several drawbacks. One of the biggest drawbacks is interference that occurs between users with different numerologies. This is essentially ICI caused by non-orthogonal subcarriers of different spacings, but in order to distinguish the term itself from a time-selective channel-induced ICI, the terms Interband Interference (IBI) or Internumerology Interference (INI) are encountered in the literature. Of the two terms used to denote this type of ICI, the latter is used in the thesis. This section explores the circumstances under which INI occurs, provides analytic expressions of INI power for both OFDM and UFMC systems, and discusses the guard bands between numerologies as a way to reduce this interference.

4.1 The Meaning of INI

Under ideal conditions and for considered waveforms (OFDM and UFMC), a single-numerology system is orthogonal. However, in multi-numerology systems the coexistence of multiple numerologies violates orthogonality, and hence causes INI [70–72]. Multiplexing users that employ different subcarrier spacings (numerologies) in the frequency domain leads to energy leakage out of the subcarrier bandwidth of one user onto the subcarriers of another user.

One such example of INI caused by a non-orthogonal system of two neighbouring users that employ different numerologies is illustrated in Fig. 4.1. In the thesis, *victim* user and *aggressor* user are used to denote the user affected by INI and the user that causes INI, respectively. From Fig. 4.1 it can also be noticed that these two users are not placed directly next to each other, but they are separated by a certain guard interval, which reduces INI but increases the system overhead at the same time.

In order to distinguish between users with different numerologies, from now on, a notation Δf_s is used to denote a user that employs a subcarrier spacing Δf within a subband s, where $s \in \{1, 2, 3, 4\}$ corresponds to $\mathcal{S}$ = {15 kHz, 30 kHz, 60 kHz, 120 kHz}, respectively.

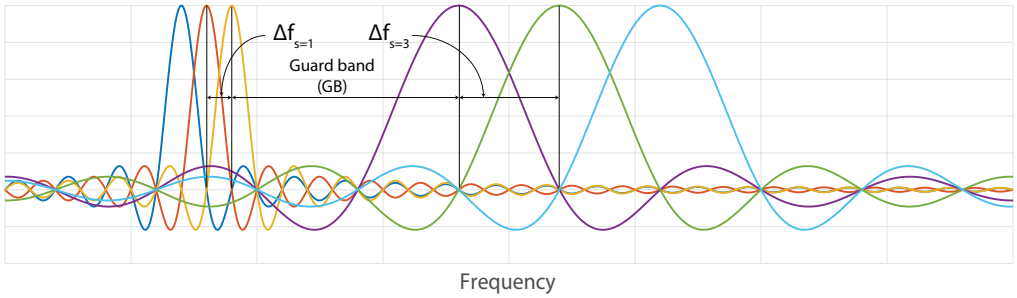


Figure 4.1: An illustration of INI between two numerologies: $\Delta f_{s=1}$ and $\Delta f_{s=3}$.

4.2 INI in OFDM Systems

This section aims to demonstrate INI in the OFDM system caused by coexistence of multiple numerologies. The closed-form expressions of INI caused by an aggressor user towards a victim user are provided and stepwise derivations are shown.

An OFDM transmission with one *victim* and one *aggressor* user is considered in the following analysis. The available bandwidth is equally shared between these two users. This further means that the first half of subcarriers within available band is occupied by the *victim* user and the second half of subcarriers is occupied by the *aggressor* user. Assuming that the total number of K_s subcarriers is available for the user with a smaller subcarrier spacing within a specific bandwidth but only the first half of them $\left(\frac{1}{2}K_s\right)$ is active[1] and assuming that this user employs a single OFDM symbol, the input signal $\mathbf{X}^s$ can be written as:

$$\mathbf{X}^s = \left[X_1^s \dots X_{\frac{K_s}{2}}^s 0 \dots 0\right]^T. \tag{4.1}$$

The transmit signal $\mathbf{x}^s$ after performing a K_s-IDFT over the entire band can be expressed as:

$$\mathbf{x}^s = \mathfrak{F}\{\mathbf{X}^s\} = \left[x_0^{s,1}, x_1^{s,1}, \dots x_{K_s-1}^{s,1}\right]^T \in \mathbb{C}^{K_s \times 1}, s \in \mathcal{S}. \tag{4.2}$$

Correspondingly, the second half of a bandwidth of total $K_{s'}$ subcarriers is occupied by the user with a larger subcarrier spacing $\left(\frac{1}{2}K_{s'}\right)$. The input signal $\mathbf{X}^{s'}$ can be written as:

$$\mathbf{X}^{s'} = \left[0 \dots 0_{\frac{K_{s'}}{2}} X_{\frac{K_{s'}}{2}+1}^s \dots X_{K_{s'}}^s\right]^T. \tag{4.3}$$

[1] This is considered as a specific example. In general, the same approach can be used to consider any arbitrary bandwidth-share.

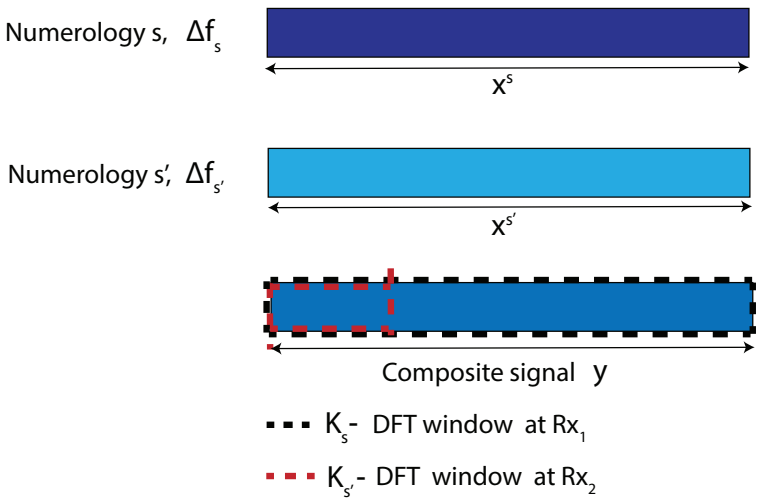


Figure 4.2: An illustration of multiplexing numerologies in both receiver scenarios.

The transmit signal after applying $K_{s'}$-IDFT over the entire band and concatenating multiple symbols of the user with a larger subcarrier spacing can be written as:

$$\begin{aligned}\mathbf{x}^{s'} = \mathfrak{F}\{\mathbf{X}^{s'}\} &= \begin{bmatrix} \mathbf{x}^{s',1} \\ \vdots \\ \mathbf{x}^{s',2^{s'-1}} \end{bmatrix} \\ &= \begin{bmatrix} \left[x_0^{s',1}, x_1^{s',1}, \dots x_{K_{s'}-1}^{s',1}\right]^T \\ \vdots \\ \left[x_0^{s',2^{s'-1}}, x_1^{s',2^{s'-1}}, \dots x_{K_{s'}-1}^{s',2^{s'-1}}\right]^T \end{bmatrix} \in \mathbb{C}^{K_s \times 1}, s' \in \mathcal{S}, s < s'.\end{aligned} \tag{4.4}$$

A larger subcarrier spacing implies shorter symbol duration, meaning that $2^{s'-1}$ shorter symbols of the user $\Delta f_{s'}$ correspond to one symbol of the user Δf_s.

In order to investigate the intrinsic interference of the system, the effects caused by the channel are neglected. Therefore, a scenario with an ideal channel and without noise is assumed. Supposing perfect synchronization between the symbols of different numerologies, as illustrated in Fig. 4.2, the composite signal of these two users is given by:

$$\mathbf{y} = \mathbf{x}^s + \mathbf{x}^{s'}, \quad \mathbf{y} \in \mathbb{C}^{K_s \times 1}. \tag{4.5}$$

In order to investigate INI, two different scenarios are considered. The first scenario considers processing at the receiver that employs a smaller subcarrier spacing (Rx_1), whereas the second scenario considers the receiver that employs a larger subcarrier spacing (Rx_2).

Let us first observe the processing at Rx_1. The user with a smaller subcarrier spacing Δf_s is considered to be the *victim* user and the user with a larger subcarrier spacing $\Delta f_{s'}$ is considered to be the *aggressor* user. After applying a K_s-DFT on the composite signal $\mathbf{y}$ as shown in Fig. 4.2, the estimated signal can be written as:

$$\hat{\mathbf{X}}^s = [\hat{X}_0^s, \hat{X}_1^s, \cdots \hat{X}_{K_s-1}^s]^T = \mathfrak{F}\{\mathbf{y}\}. \tag{4.6}$$

A single estimated symbol at subcarrier l is given by:

$$\hat{X}_l^s = \hat{X}_{\mathrm{des}_l}^s + \hat{X}_{\mathrm{INI}_l}^s, \;\; l \in \left[0, \frac{K_s}{2}\right] \tag{4.7}$$

where $\hat{X}_{\mathrm{des}_l}^s$ is the desired part and $\hat{X}_{\mathrm{INI}_l}^s$ is interference caused by the *aggressor* user within a subband s. After the derivation provided in Appendix 8, the INI power caused by a single OFDM symbol of the *aggressor* user towards the *victim* user is given by:

$$\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l,1}} = \frac{1}{K_{s'}K_s} \sum_{k=0}^{\frac{1}{2}K_{s'}-1} \underbrace{\mathbb{E}\left[\left|X^{s',1}_{\left(\frac{1}{2}K_{s'}+k\right)}\right|\right]^2}_{=1} \frac{\sin^2\left[\pi\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}{\sin^2\left[\frac{\pi}{K_{s'}}\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}. \tag{4.8}$$

Taking into account all $2^{s'-1}$ OFDM symbols of the user $\Delta f_{s'}$, with equal interference powers of all symbols, the INI power that affects the user Δf_s, $\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l}}$, can be expressed as:

$$\begin{aligned}\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l}} &= 2^{s'-1}\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l,1}} \\ &= \frac{1}{K_{s'}^2} \sum_{k=0}^{\frac{1}{2}K_{s'}-1} \frac{\sin^2\left[\pi\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}{\sin^2\left[\frac{\pi}{K_{s'}}\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}.\end{aligned} \tag{4.9}$$

The spectral distance is represented by $k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}$ between the interfering subcarrier at $k + \frac{1}{2}K_s$ of the *aggressor* user and the subcarrier l of the *victim* user. According to (4.9), it can be observed that INI power depends on the applied numerologies, i.e., on the scalability factor $2^{s'-1}$.

The INI power at Rx_1 is shown in Fig. 4.3, where two neighbouring users with different numerologies are represented in terms of their spectrum. On the left-hand side, the *victim* user $\Delta f_{s=1}$ is represented by blue colour. On the right-hand side, the *aggressor* user $\Delta f_{s' \in \{2,3,4\}}$ is represented by red, green and magenta colours, respectively. It can be observed that the edge subcarriers experience higher INI compared to the middle subcarriers due to the high spectrum side lobes of the neighbouring numerology. Moving away from the edges, the side lobes of the spectrum are more suppressed, and therefore INI is lower. It is visible that the interference is zero whenever 15 kHz subcarriers fall onto a multiple of 30 kHz, 60 kHz and 120 kHz, respectively.

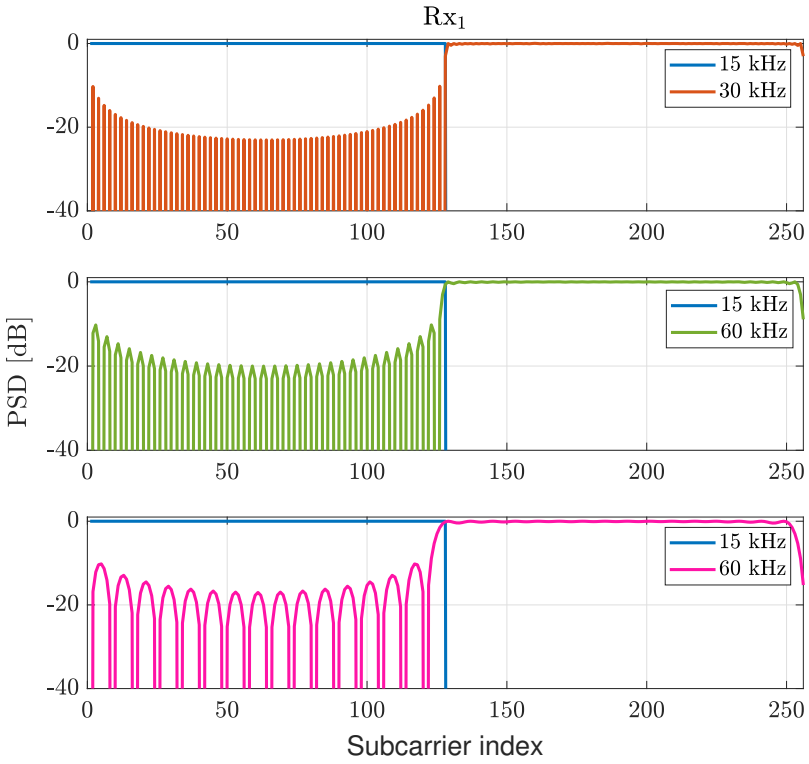


Figure 4.3: INI power on the user Δf_s caused by the user $\Delta f_{s'}$.

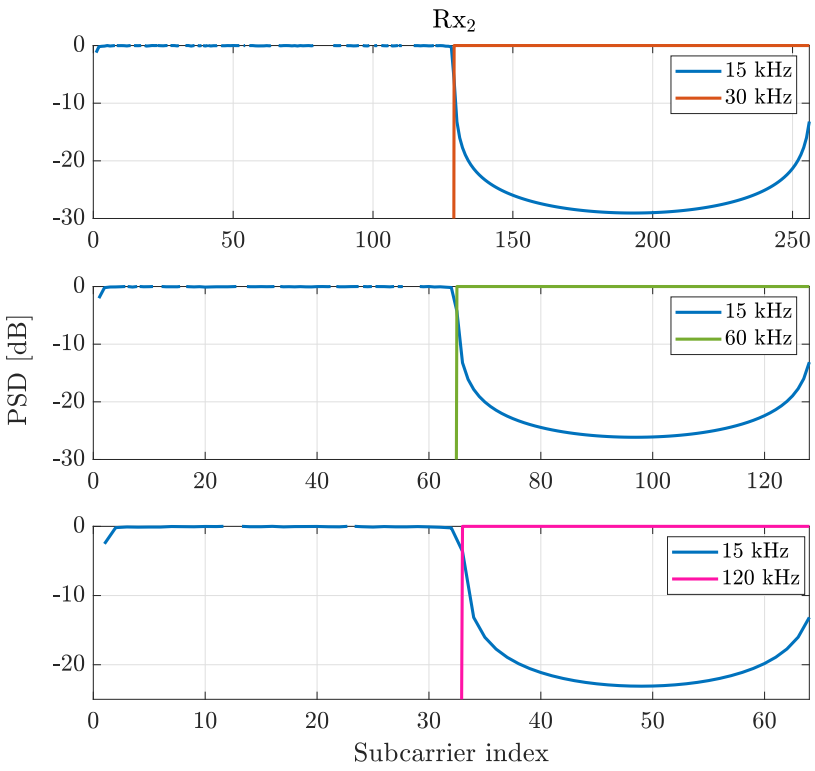


Figure 4.4: INI power on the user $\Delta f_{s'}$ caused by the user Δf_s.

At Rx_2, the user that employs a larger subcarrier spacing becomes the *victim* user and the *aggressor* is the one that employs a smaller subcarrier spacing. Here, a $K_{s'}$-DFT is applied only to $1/2^{s'-1}$ of the samples transmitted by Δf_s user as shown in Fig. 4.2, taking into account that symbol duration is scaled according to (2.17). Utilizing the vector-valued input-output relationship, the INI power $\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_{s'}}}$ caused by the user Δf_s towards the user $\Delta f_{s'}$ can be expressed as:

$$\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_{s'}}} = \mathrm{diag}\Big(\mathbf{D}^{s'}\mathbf{D}^{s,H}_{\mathrm{sub}}\underbrace{\mathbb{E}\Big[\mathbf{X}^s\mathbf{X}^{s,H}\Big]}_{\mathbf{I}_{K_{s'}}}\mathbf{D}^{s}_{\mathrm{sub}}\mathbf{D}^{s',H}\Big), \tag{4.10}$$

where $\mathbf{D}^{s'}$ is a square DFT matrix of dimension $K_{s'}$ and $\mathbf{I}$ is an identity matrix. It is assumed that the user that employs a smaller subcarrier spacing occupies a lower part of the bandwidth; therefore, as shown in (4.1) the first $K_{s'}$ subcarriers of vector $\mathbf{X}^s \in \mathbb{C}^{K_s \times 1}$ are active and the remaining subcarriers are set to zeros. Correspondingly, $\mathbf{D}^s_{\mathrm{sub}}$ is a square sub-matrix representing the first $K_{s'}$ samples of a DFT matrix of dimension K_s.

Fig. 4.4 shows that each subcarrier is affected by INI at Rx_2. It can be observed that the spectrum

roll-off becomes steeper as its subcarrier spacing becomes smaller. In this specific case, the smallest INI power is achieved by employing 30 kHz subcarrier spacing. However, edge subcarriers for all cases are almost equally affected.

4.3 INI in UFMC Systems

Similarly to the OFDM case, the INI power in the UFMC system is derived for two scenarios:

1. When interference coming from a user with larger subcarrier spacing.
2. When interference coming from a user with smaller subcarrier spacing.

Considering the processing at Rx_1, the INI power in the UFMC system is given by:

$$\sigma^{2,\mathrm{ufmc}}_{\mathrm{INI}_{\Delta f_s},l} = \sum_{k=0}^{\frac{1}{2}K_{s'}-1} |f_l|^2 \frac{\sin^2\left[\pi\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}{\sin^2\left[\frac{\pi}{K_{s'}}\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}, \tag{4.11}$$

where

$$f_l = f_H\Big[< \mathrm{v} - f_{\mathrm{ce}_k} >_{K_s}\Big], \;\; \mathrm{v} \in \Big[0, K_s - 1\Big], f_{\mathrm{ce}_k} \in \mathbb{N}, \tag{4.12}$$

and l = v at $\left[0, \frac{K_s}{2} - 1\right]$. A circular shift of the subcarrier v by the block center frequency f_{ce_k} is denoted by $< \mathrm{v} - f_{\mathrm{ce}_k} >_{K_s}$, and the modulo operation is denoted by $<>$ (($\mathrm{v} - f_{\mathrm{ce}_k}$) modulo K_s). The frequency domain of the Hanning window at subcarrier v can be written as:

$$\begin{aligned} f_H[\mathrm{v}] = {} & \frac{1}{2}\exp^{-i\pi \mathrm{v}\frac{L_{F_{s'}}-1}{K_s}} \frac{\sin\left[\frac{\pi \mathrm{v} L_{F_{s'}}}{K_s}\right]}{\sin\left[\frac{\pi \mathrm{v}}{K_s}\right]} - \frac{1}{4}\exp^{-i\pi\left(<\mathrm{v}-\frac{K_s}{L_{F_{s'-1}}}>_{K_s}\right)\frac{L_{F_{s'}}-1}{K_s}} \frac{\sin\left[\frac{\pi\left(<\mathrm{v}-\frac{K_s}{L_{F_{s'-1}}}>_{K_s}\right)L_{F_{s'}}}{K_s}\right]}{\sin\left[\frac{\pi\left(<\mathrm{v}-\frac{K_s}{L_{F_{s'-1}}}>_{K_s}\right)}{K_s}\right]} \\ & - \frac{1}{4}\exp^{-i\pi\left(<\mathrm{v}-K_s+\frac{K_s}{L_{F_{s'-1}}}>_{K_s}\right)\frac{L_{F_{s'}}-1}{K_s}} \frac{\sin\left[\frac{\pi\left(<\mathrm{v}-K_s+\frac{K_s}{L_{F_{s'-1}}}>_{K_s}\right)L_{F_{s'}}}{K_s}\right]}{\sin\left[\frac{\pi\left(<\mathrm{v}-K_s+\frac{K_s}{L_{F_{s'-1}}}>_{K_s}\right)}{K_s}\right]}. \end{aligned} \tag{4.13}$$

As observed from (4.11), the closed-form expression is composed of two parts - a pure OFDM part obtained according to (4.9) and a filter transfer function $|f_l|^2$. It is a frequency transform of the Hanning window adapted to each of the UFMC subbands. Term (4.11) is normalized with respect to the maximum value.

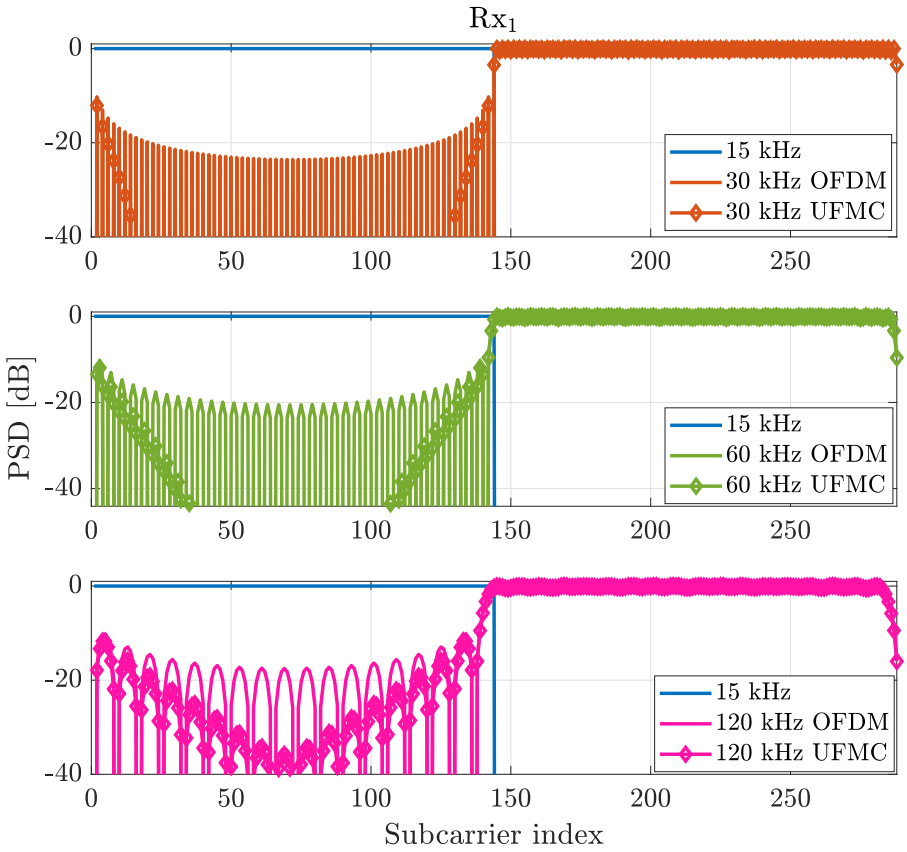


Figure 4.5: PSD of OFDM and UFMC for three numerologies:$\Delta f_{s'=2}$, $\Delta f_{s'=3}$ and $\Delta f_{s'=4}$.

The UFMC technique uses filtering which smoothes the rectangular pulse of the OFDM signal, thereby providing better spectrum confinement and steeper spectrum roll-off. In Fig. 4.5 three different scenarios with $\Delta f_{s=1}$ user on one side and $\Delta f_{s'\in\{2,3,4\}}$ user on the other side are shown. In all these cases, the filter length employed by numerology Δf_s, L_{F_s}, is proportionally scaled according to the factor $2^{s'-1}$, i.e., a longer filter length corresponds to a smaller subcarrier spacing, and vice versa. Interestingly, the edge subcarriers are almost equally affected by INI as in OFDM, but moving away from the edges the filter abates the spectrum side lobes, and hence the INI power.

At Rx_2, in a similar way as in the OFDM case, the INI power $\sigma^{2,\mathrm{ufmc}}_{\mathrm{INI}_{\Delta f_{s'}}}$ is expressed using a vector-valued input-output relationship:

$$\sigma^{2,\mathrm{ufmc}}_{\mathrm{INI}_{\Delta f_{s'}}} = \Big|\bar{\mathbf{D}}_{s'} \; \boldsymbol{\Xi} \; \underbrace{\sum_{b} \mathbf{F}_b \; \mathbf{D}_b^H \; \mathbf{X}_b^s}_{transmit \quad signal}\Big|^2 \in \mathbb{C}^{K_{s'}\times 1}, \tag{4.14}$$

where $\mathbf{X}_b^s \in \mathbb{C}^{K_s \times 1}$ is the input signal of a specific subband b defined in (2.8). Matrix $\bar{\mathbf{D}}_{s'}$ is given by:

$$\bar{\mathbf{D}}_{s'} = [\mathbf{D}_{K_{s'} \times K_{s'}} \mathbf{0}_{K_{s'} \times K_{s'}}]. \tag{4.15}$$

Filter matrix $\mathbf{F}_b \in \mathbb{C}^{(K_s + L_{\mathrm{F}_s}) \times K_s}$ is a Toeplitz matrix with first column $[\mathbf{f}_b \ \mathbf{0}_{1 \times K_s}]^T$ and the first row $[f_b(0) \ \mathbf{0}_{1 \times K_s - 1}]$, where $\mathbf{f}_b$ is defined in (2.10). In order to have the same receiver complexity as in CP-OFDM, instead of a $2K_{s'}$-DFT only a $K_{s'}$-DFT is executed as defined by a matrix $\bar{\mathbf{D}}_{s'}$. To recover the data successfully by applying a different receiver, a cyclic property of the system has to be preserved as explained in [51]. For that propose, matrix $\boldsymbol{\Xi}$ is defined as:

$$\boldsymbol{\Xi} = \begin{bmatrix} \mathbf{I}_{K_{s'}} \ [\mathbf{I}_{L_{\mathrm{F}_{s'}}} \ \mathbf{0}^T_{(K_{s'} - L_{\mathrm{F}_{s'}}) \times L_{\mathrm{F}_{s'}}}]^T \ \mathbf{0}_{K_{s'} \times (K_s + L_{\mathrm{F}_{s'}} - K_{s'} - L_{\mathrm{F}_{s'}})} \\ \mathbf{0}_{K_{s'} \times (K_s + L_{\mathrm{F}_{s'}} - K_{s'} - L_{\mathrm{F}_{s'}})} \ \mathbf{I}_{K_{s'}} \ [\mathbf{I}_{L_{\mathrm{F}_{s'}}} \ \mathbf{0}^T_{(K_{s'} - L_{\mathrm{F}_{s'}}) \times L_{\mathrm{F}_{s'}}}]^T \end{bmatrix}. \tag{4.16}$$

4.4 Simulation Results

Fig. 4.6 provides a verification of the analytically derived INI power in (4.9) and (4.11), when the user with a larger subcarrier spacing acts as the *aggressor* user. In this specific case, two adjacent numerologies with equally shared bandwidth are considered. The *victim* user employs 15 kHz subcarrier spacing, i.e., $\Delta f_{s=1}$, whereas the *aggressor* user employs 30 kHz subcarrier spacing, i.e., $\Delta f_{s=2}$. Considering only one RB of 12 subcarriers for the *victim* user and 6 subcarriers for the *aggressor* user, Fig. 4.6 shows that the analytical solution coincides with the simulation results for all filter lengths as well as for the OFDM case. In these simulations, the Hanning filter is employed. It can be observed that longer and thus sharper filters decrease the INI power. However, as previously mentioned, a long filter introduces intrinsic interference within doubly-selective channels. Therefore, there is a tradeoff between spectral efficiency and intrinsic interference which results in the optimal filter length [73].

In addition to the optimal choice of filters, the INI power can also be controlled by applying a guard band between adjacent numerologies as shown in Fig. 4.1. This implies the separation of numerologies, which, on the one hand, mitigates the most influential effect of INI, but which certainly reduces the frequency efficiency of the system on the other hand. The effects of the filters as well as the guard bands on the cumulative INI power of all affected subcarriers within a corresponding subband and for both receiver scenarios are presented in Fig. 4.7 and Fig. 4.8. Fig. 4.7 considers a victim user that employs 15 kHz subcarrier spacing and an aggressor user that employs 30 kHz, 60 kHz or 120 kHz. Fig. 4.8 considers an aggressor user that employs 15 kHz subcarrier spacing and a victim user that employs 30 kHz, 60 kHz or 120 kHz. In both cases INI has a similar behaviour, that is, the INI power achieves its maximum with zero guard subcarriers and a short filter length, whereas, with increasing filter length and number of guard subcarriers, INI decays. Despite the fact that INI is not present on each subcarrier in the case of Rx_1, the total interference power is greater than in

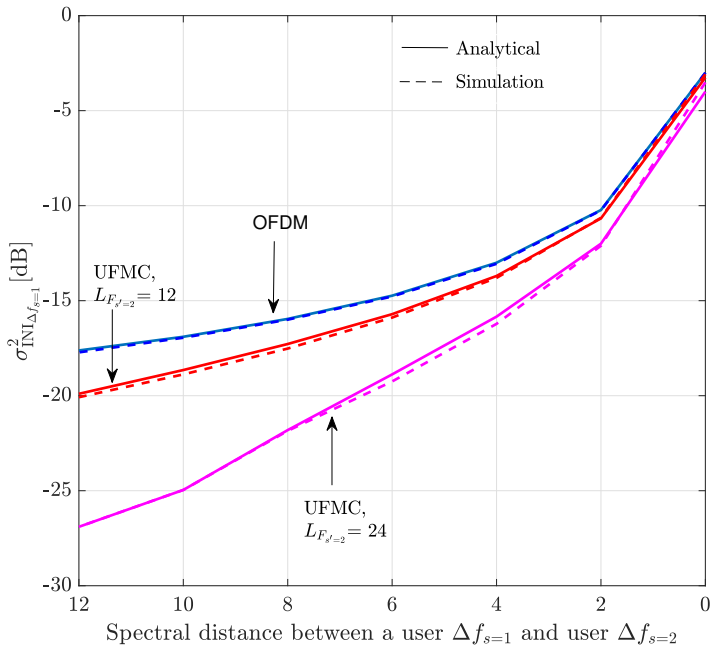


Figure 4.6: INI power at Rx_1 - analytical vs. simulation results.

the case of Rx_2, at least for shorter filters. Furthermore, smaller subcarrier spacings have steeper spectrum roll-off, which directly implies a lower level of INI.

There is also a tradeoff between using longer filters or assuming several guard subcarriers in terms of time-frequency efficiency; a longer filter causes the loss of time efficiency due to the CP/ZP, while guard subcarriers are harmful to the frequency efficiency. The time-frequency efficiency can be defined as:

$$\begin{aligned} \mathrm{TF} &= \frac{B_{\mathrm{u}}}{B}\frac{N_{\mathrm{u}}}{N}, \\ &= \frac{K-G}{K}\frac{K}{K+L_F+N_{ZP}-1}, \end{aligned} \tag{4.17}$$

where B is the total bandwidth of 8.64 MHz that corresponds to K = 576 subcarriers of 15 kHz subcarrier spacing, and B_u is the useful bandwidth reduced by a number of guard subcarriers G; N_u is the number of useful time domain samples that is equal to K and N is the entire symbol length including the filter and ZP length.

In Fig. 4.9 the joint effect of the time-frequency efficiencies and achievable rates, in terms of different filter lengths and different number of guard subcarriers, are compared. The metric C_e shown on the vertical axis is the product of the mean achievable rate defined as $\mathrm{R}_{\Delta f_s} = \log_2(1+\bar{\gamma}_{\Delta f_s})$ over

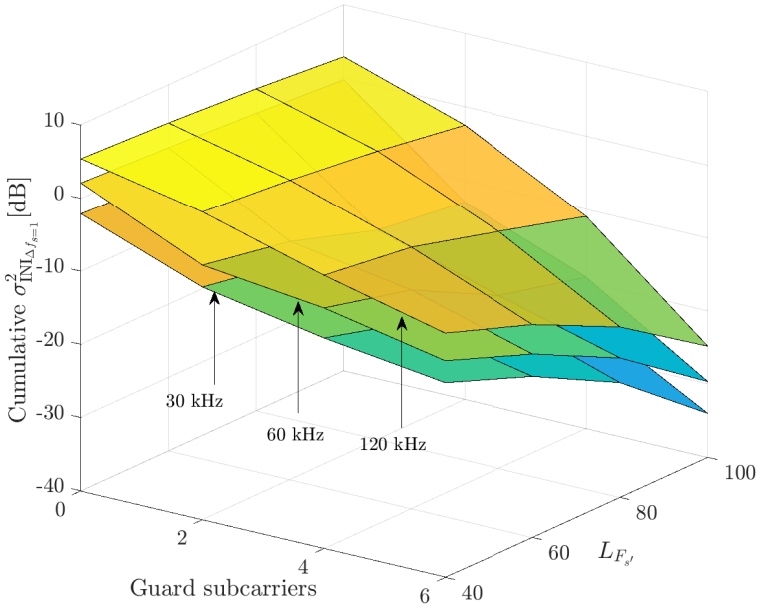


Figure 4.7: INI power at Rx_1 - a victim user employs 15 kHz subcarrier spacing and an aggressor user employs 30 kHz, 60 kHz or 120 kHz.

all used subcarriers and the time-frequency efficiency TF, and it is normalized with respect to the maximum achieved value:

$$C_e = \mathrm{R}_{\Delta f_s}\mathrm{TF}. \tag{4.18}$$

Two different regimes are compared: the first regime considers a low noise variance ($\sigma_{\mathrm{n}}^2 = 10^{-8}$) compared to the INI power and the second regime considers a higher noise variance or more precisely the noise power in the order of the INI power ($\sigma_{\mathrm{n}}^2 = 10^{-1}$). In the former case, the performance is determined by the tradeoff between TF efficiency and the residual INI power. It can be observed that the maximum is achieved for the largest guard band and long filter length L_{F_s} = 100. In the latter case when the noise power is similar to the INI power, the maximum is achieved for the shortest considered filter length and two guard subcarriers. However, it can be noticed that in both cases the impact of the guard band is smaller compared to the impact of filter lengths.

4.5 Spectral Emission Mask

One of the challenges regarding multi-numerology systems is to find an adaptive guard band between numerologies. One way to do this is to define a spectral mask, i.e., to meet OOB emission

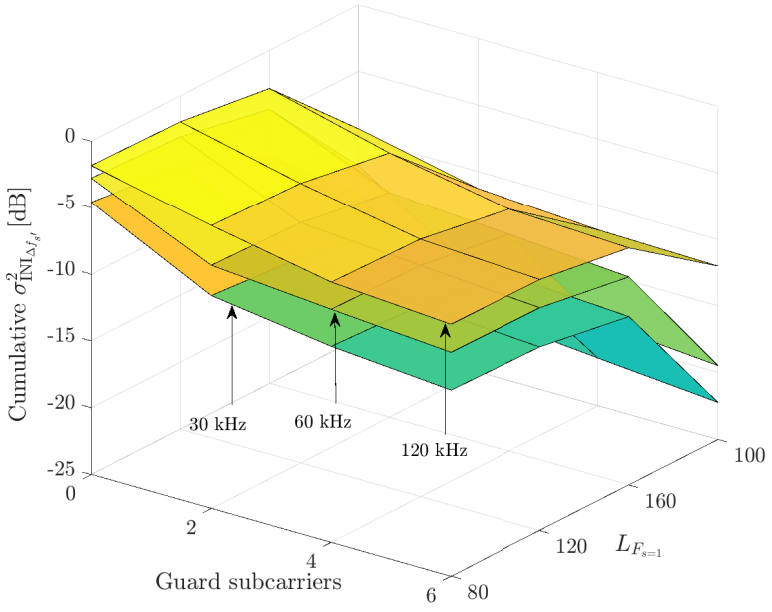


Figure 4.8: INI power at Rx_2 - an aggressor user employs 15 kHz subcarrier spacing and a victim user employs 30 kHz, 60 kHz or 120 kHz.

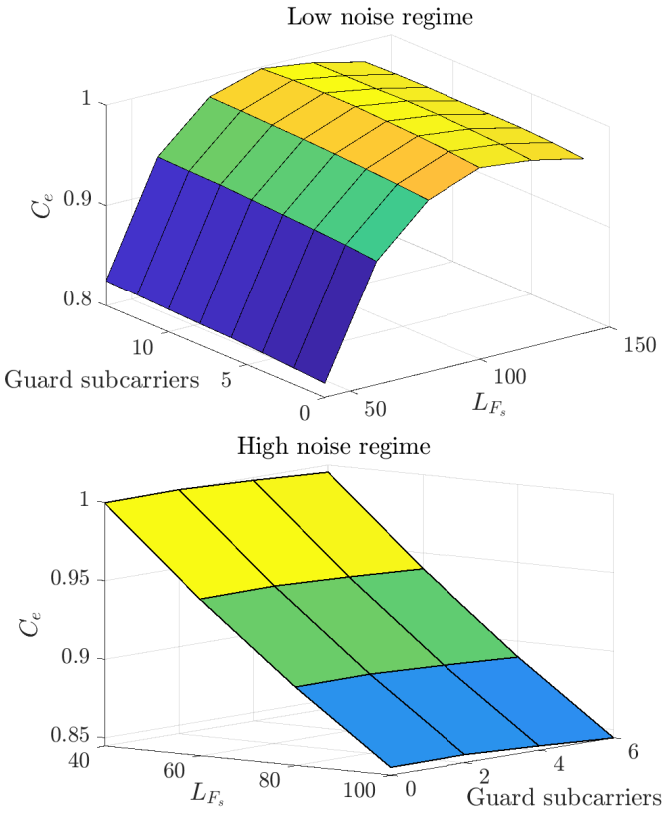


Figure 4.9: Joint effect of the user achievable rate and time-frequency efficiency w.r.t. the guard bands and filter lengths.

requirements. The Adjacent Channel Leakage power Ratio (ACLR) and Operating Band Unwanted Emissions (OBUE) methods are most commonly used in the literature [74].

In contrast to these methods which propose more complex mask structures, a simplified form of guard band definition is considered in the thesis. Namely, a certain threshold $\sigma^2_{\mathrm{INI}^{\max}}$ is determined, representing an upper bound on the allowed INI power, as shown in Fig. 4.10. Defining this threshold, the number of guard subcarriers between certain numerologies is defined in order to mitigate the INI that lies above this threshold. The selection of the threshold value is a tradeoff between the guard band overhead and INI. That is, smaller $\sigma^2_{\mathrm{INI}^{\max}}$ implies larger guard band overhead, and therefore reduced time-frequency efficiency. Yet, less severe INI is present with more guard subcarriers. This threshold is chosen such that a dominant effect compared to other interference and noise sources is prevented, on the one hand, and on the other hand, the strongest INI is mitigated with a defined guard band, keeping the spectral efficiency high. Since the size of the spectrum roll-off factor depends on numerology, a different number of guard subcarriers between adjacent numerologies is obtained for the fixed threshold. Reasonably, a larger scale factor between numerologies yields more guard subcarriers, and vice versa.

In Chapter 3 the SINR expression is provided based on ICI and ISI as a result of doubly-selective channels and the channel estimation error due to imperfect channel knowledge. Additionally, the final expression for SINR has to include INI as a consequence of the multi-numerology system as well. However, in the further analysis only $\sigma^2_{\mathrm{INI}^{\max}}$ will be taken into account which simplifies the solution of the optimization problems formulated in Section 5:

$$\bar{\gamma}_{\Delta f} \geq \frac{\sigma^2_{\mathrm{d}}}{\sigma^2_{\mathrm{n}} + \sigma^2_{\mathrm{ICI}_{\Delta f}} + \sigma^2_{\mathrm{ISI}_{\Delta f}} + \sigma^2_{\mathrm{e}_{\Delta f}} + \sigma^2_{\mathrm{INI}^{\max}_{\Delta f}}}. \tag{4.19}$$

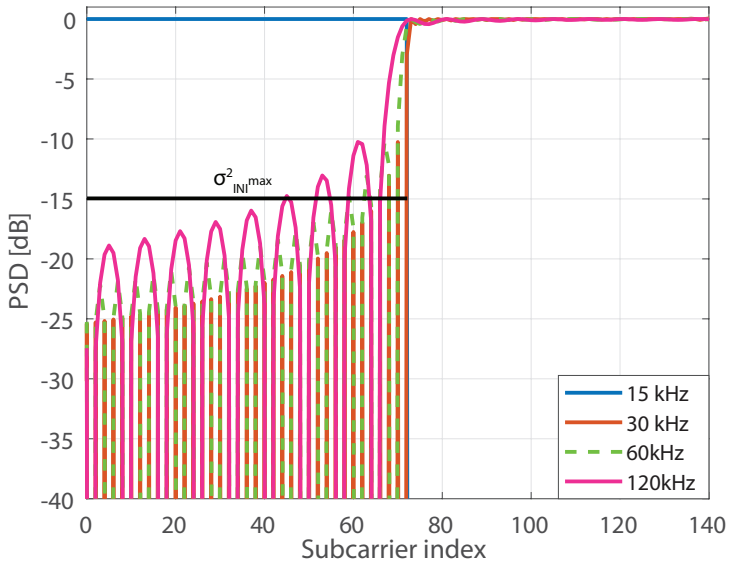


Figure 4.10: An example of the threshold $\sigma^2_{\mathrm{INI}^{\max}}$.

Chapter 5

Optimal Resource Allocation based on Mixed Numerology

In this chapter, two main approaches for optimal resource and numerology allocation in a multi-user scenario are proposed. A novel optimization strategy, which enables optimal allocation of bandwidth amongst users with different numerologies, is considered within the first approach. An optimization method for the joint resource and numerology allocation is proposed within the second approach. The methods are based on frequency multiplexing of numerologies, and time multiplexing is not considered.

The optimal choice of resources and numerology depends on the PHY parameters used to model the long-term average SINR defined in (4.19). The achievable rate $\mathrm{R}_{\Delta f_s}$, as a fundamental metric is exploited in our optimization formulation. Given the SINR $\bar{\gamma}_{\Delta f_s}$ in a subband with subcarrier spacing Δf_s, it is defined as:

$$\mathrm{R}_{\Delta f_s} = \log_2(1 + \bar{\gamma}_{\Delta f_s}). \tag{5.1}$$

5.1 Optimal Resource Allocation with Predefined Numerology

In this scenario, it is assumed that the user devices are manufactured to operate with fixed subcarrier spacings. For example, vehicles may use larger subcarrier spacings and mobile phones may use 15 kHz. The idea of this section is to provide an insight into how different numerologies adopt to the same PHY conditions. In this regard, the optimal resource allocation is investigated amongst users with different numerologies under the same channel conditions. One example of a bandwidth partition is shown in Fig. 5.1. The bandwidth is utilized by four users that use different numerologies from set $\mathcal{S} = \{15, 30, 60, 120\}$ kHz and belong to appropriate subbands, $\Delta f_{s=1}$, $\Delta f_{s=2}$, $\Delta f_{s=3}$ and $\Delta f_{s=4}$, respectively[1]. Since users employ different numerologies they experience different ICI and ISI and therefore different SINR values. The resource allocation exclusively depends on the SINR that each user achieves.

[1] Notice that one subband corresponds to one user.

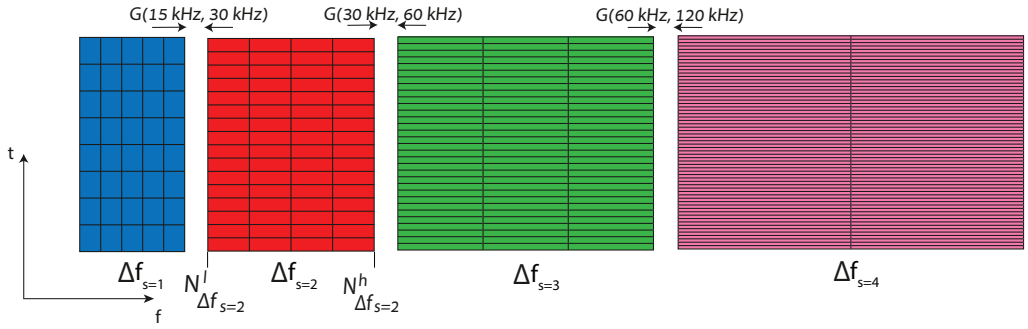


Figure 5.1: An example of resource allocation of four users with different numerologies.

The goal of the optimization is to serve all users equally under the same channel conditions. This can be achieved by maximizing the minimum achievable rate of each user. Hence, the optimization problem is defined as:

$$
\begin{aligned}
\max_{N^h_{\Delta f_s}, N^l_{\Delta f_s}, c_s} \quad & z \\
\text{subject to} \quad & \log_2(1+\bar{\gamma}_{\Delta f_s})(N^h_{\Delta f_s} - N^l_{\Delta f_s}) \geq z, \;\; \Delta f_s \in \mathcal{S}, z \in \mathbb{R} \\
& N^l_{\Delta f_s} \leq N^h_{\Delta f_s}, \;\; \{N^l_{\Delta f_s}, N^h_{\Delta f_s}\} \in \mathbb{N}_0 \\
& N^h_{\Delta f_{s-1}} \leq N^l_{\Delta f_s}, \\
& N^h_{\Delta f_s} - N^l_{\Delta f_s} = 2^{s-1} c_s, \;\; s \in \{1,2,3,4\}, c_s \in \mathbb{N} \\
& N^l_{\Delta f_{s+1}} - N^h_{\Delta f_s} \geq G_{\Delta f_s, \Delta f_{s+1}}, \\
& \sum_{s \in \mathcal{S}} (N^h_{\Delta f_s} - N^l_{\Delta f_s} + G_{\Delta f_s, \Delta f_{s+1}}) = K^{\max}_{15\,\text{kHz}},
\end{aligned}
\tag{5.2}
$$

where z is a lower bound on the achievable rate, Δf_s denotes an appropriate subcarrier spacing and $N^h_{\Delta f_s}$ and $N^l_{\Delta f_s}$ are upper and lower bounds of the corresponding subband, respectively, as illustrated in Fig. 5.1. More precisely, the borders of the subbands are optimized according to the SINR, equalizing the achievable rates of the users.

The first constraint ensures that the lower border of each subband lies below the upper border:

$$
N^l_{\Delta f_s} \leq N^h_{\Delta f_s}, \qquad \Delta f_s \in \mathcal{S}, \tag{5.3}
$$

within set $\mathcal{S}$ of all subbands.

The users are sorted in an incremental order, so that the user with the smallest subcarrier spacing corresponds to the first subband, and the user with the largest subcarrier spacing corresponds to the

last subband. This is represented by the constraint:

$$N^h_{\Delta f_{s-1}} \leq N^l_{\Delta f_s}. \tag{5.4}$$

An inevitable constraint in the optimization is related to scalable numerology with respect to the basic 15 kHz subcarrier spacing. It enforces the number of allocated subcarriers to be a multiple of the power of two:

$$N^h_{\Delta f_s} - N^l_{\Delta f_s} = 2^{s-1}c_s, \;\; c_s \in \mathbb{N}, \tag{5.5}$$

where c_s is a non-negative integer value. By applying this constraint it is ensured that each subband is occupied by a certain user, i.e., each subband is active. The minimum number of subcarriers that is accommodated within a certain subband is one for the 15 kHz user, two for the 30 kHz user, four for the 60 kHz user and eight for the 120 kHz user.

Having multiple active numerologies, it is necessary to define a guard band between them in order to reduce INI. As already mentioned in Chapter 4, a guard band between users is predefined with respect to $\sigma^2_{\mathrm{INI}^{\max}}$. To account for the guard band, the following constraint is defined:

$$N^l_{\Delta f_{s+1}} - N^h_{\Delta f_s} \geq G_{\Delta f_s, \Delta f_{s+1}}, \tag{5.6}$$

meaning that the guard band $G_{\Delta f_s, \Delta f_{s+1}}$ is the difference between the lower bound of the following user who occupies Δf_{s+1} and the upper bound of the previous user who occupies Δf_s. The subcarriers assigned to the guard band are given with respect to the minimum 15 kHz subcarrier spacing.

In order to utilize all available resources, all subcarriers are allocated, either to some of the users or to the guard band:

$$\sum_{s \in \mathcal{S}} N^h_{\Delta f_s} - N^l_{\Delta f_s} + G_{\Delta f_s, \Delta f_{s+1}} = K^{\max}_{15\,\mathrm{kHz}}. \tag{5.7}$$

The total number of subcarriers occupied by all considered users within a certain bandwidth is defined with respect to the 15 kHz subcarrier spacing and denoted as $K^{\max}_{15\,\mathrm{kHz}}$.

The optimization problem proposed in (5.2) is given as an ILP. Despite the fact that this is a combinatorial problem, it is relatively easy to solve due to the small number of optimization variables.

5.1.1 Simulation Results

By solving problem (5.2), the optimal resource allocations for the purpose of equalization of users' achievable rates in two different scenarios are shown in Fig. 5.2 and Fig. 5.3. The list of parameters used for the simulations is summarized in Table 5.1.

The first scenario, representing a TDL-A channel model with ξ = 80 ns of RMS delay spread and various velocities from set {5, 50, 100, 200, 300} km/h, is shown in Fig. 5.2. The bandwidth of

Table 5.1: Simulation parameters

parameter	**value**			
Subcarrier spacing [kHz]	15	30	60	120
Number of symbols per subframe	14	28	56	112
CP duration [µs]	4.69	2.34	1.17	0.59
Velocity [km/h]	$\{5, 50, 100, 200, 300\}$			
Doppler spread [Hz]	$\{27.3, 273.3, 546.7, 1093.3, 1640\}$			
Bandwidth [MHz]	5.76			
Carrier frequency [GHz]	5.9			
Channel model & ξ [ns]	TDL-A {80, 380}			
Pilot pattern	Diamond			

5.76 MHz is shared amongst four users with different numerologies from set $\mathcal{S}$. Additionally, some resources are utilized by the guard band due to INI. As already emphasized, the resource allocation is based on the channel conditions and INI as a consequence of mixed numerology. With small RMS delay spread and low velocities almost an equal resource allocation can be noticed amongst users that employ different numerologies. When the velocity is increased, the user that employs 15 kHz tends to occupy more resources in order to equalize the achievable rate with the other users, whereas the amount of resources occupied by 60 kHz and 120 kHz users decreases. This is a consequence of improved robustness of a larger subcarrier spacing to ICI. With larger velocities ICI becomes more dominant over other parameters. The 30 kHz subcarrier spacing remains quite stable, with a tendency to reduce the number of resources at higher velocities.

In the second scenario shown in Fig. 5.3, for ξ = 380 ns RMS delay spread, an additional ISI becomes influential. As emphasized before, large subcarrier spacings employ small duration of symbols as well as CP, and they are less robust against ISI. Additionally, the channel estimation error is greater with larger subcarrier spacings due to a large interpolation error. Hence, already in the example with low velocity it can be observed that the 120 kHz user allocates more subcarriers than the others. Similarly to the previous example, the need for more bandwidth decreases when increasing velocity.

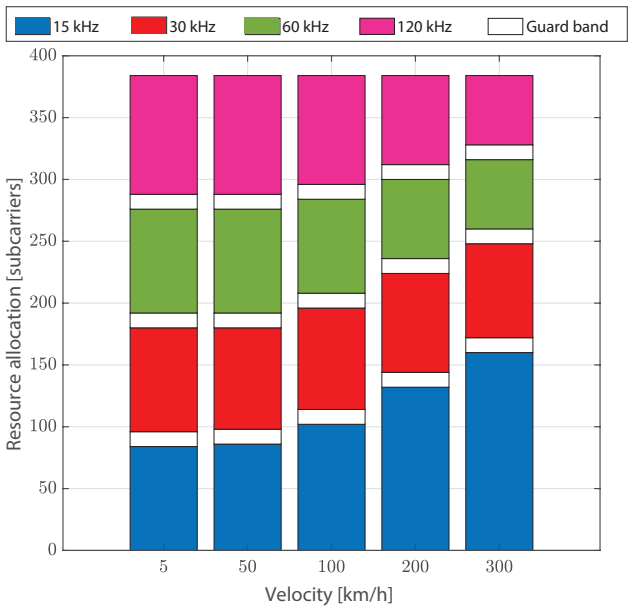


Figure 5.2: Resource allocation when $\xi = 80$ ns.

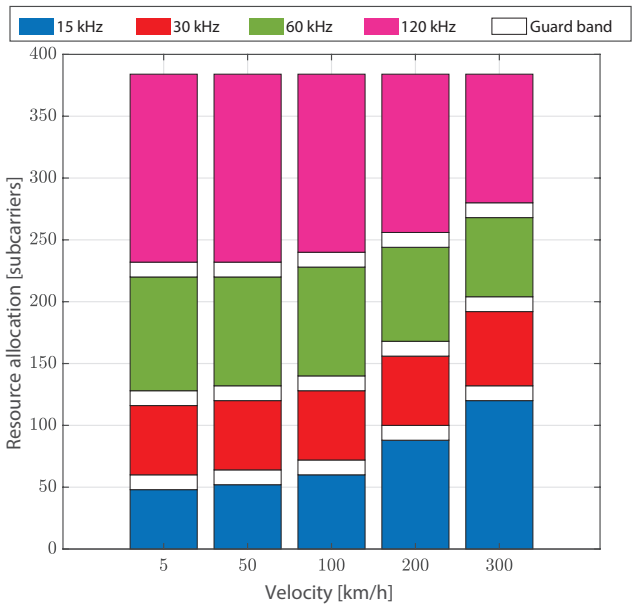


Figure 5.3: Resource allocation when $\xi = 380$ ns.

5.2 Optimal Resource and Numerology Allocation

In this section, an algorithm for the joint resource and numerology allocation in a multi-user mixed-numerology scenario is investigated. Similarly to Section 5.1, the base station assigns optimal resources and additionally an optimal numerology for transmission to each of the users, based on the channel conditions they experience. The optimization is initially given as an ILP. To reduce the computational complexity of the optimization, a linear relaxation (Linear Programming (LP) method) of the problem is also proposed. The performance of the proposed methods are investigated by numerical simulations and compared to an LTE-compliant scenario with fixed 15 kHz subcarrier spacing as well as heuristic approaches.

5.2.1 Resource and Numerology Allocation obtained by applying the ILP Algorithm

A realistic scenario where users experience different channel conditions is assumed. In such a situation the numerology is adapted to each user, with the goal of achieving the same rate performances. The minimum achievable rates of all users is maximized. An optimization problem is formulated as an ILP and solved as:

$$
\begin{aligned}
&\text{maximize} \quad z \\
&\text{subject to} \quad \sum_{\Delta f_s \in \mathcal{S}} \log_2(1+\bar{\gamma}^u_{\Delta f_s}) \Delta N^u_{\Delta f_s} \geq z, \;\; \Delta f_s \in \mathcal{S}, u \in \mathcal{U}, z \in \mathbb{R} \\
&\qquad N^{u,l}_{\Delta f_s} \leq N^{u,h}_{\Delta f_s}, \;\; \forall u \in \mathcal{U}, \;\; \{N^{u,l}_{\Delta f_s}, N^{u,h}_{\Delta f_s}\} \in \mathbb{N}_0 \\
&\qquad \sum_{\Delta f_s \in \mathcal{S}} a^u_{\Delta f_s} = 1, \;\; a^u_{\Delta f_s} \in \{0,1\} \\
&\qquad a^u_{\Delta f_s} \leq \Delta N^u_{\Delta f_s} \leq K^{\max}_{\Delta f_s} a^u_{\Delta f_s}, \;\; a^u_{\Delta f_s} \in \{0,1\} \\
&\qquad N^{u',l}_{\Delta f_{s'}} - N^{u,h}_{\Delta f_s} \geq G_{\Delta f_s, \Delta f_{s'}} a^u_{\Delta f_s}, \;\; \Delta f_{s'} > \Delta f_s, u' \neq u \\
&\qquad N^{u+1,l}_{\Delta f_s} - N^{u,h}_{\Delta f_s} = 0, \\
&\qquad \Delta N^u_{\Delta f_s} = 2^{s-1} c_s, \;\; s \in \{1,2,3,4\}, c_s \in \mathbb{N}_0,
\end{aligned}
\tag{5.8}
$$

where $\Delta N^u_{\Delta f_s}$ denotes the number of subcarriers allocated to user u in the subband Δf_s. By assuming that consecutive subcarriers are allocated to the user, the number of allocated subcarriers can be calculated as a difference between the highest and the lowest subcarrier allocated to the user u:

$$
\Delta N^u_{\Delta f_s} = N^{u,h}_{\Delta f_s} - N^{u,l}_{\Delta f_s}. \tag{5.9}
$$

and

$$N_{\Delta f_s}^{u,l} \leq N_{\Delta f_s}^{u,h}. \tag{5.10}$$

The restriction that each user can occupy only subcarriers from a single subband is imposed. This, however, does not impact the optimality of the problem formulation, since each user has a unique preferred subcarrier spacing. This is achieved by introducing a binary indicator variable $a_{\Delta f}^{u}$ that specifies whether or not user u is active within the subband Δf_s. The constraint:

$$\sum_{\Delta f_s \in \mathcal{S}} a_{\Delta f_s}^{u} = 1, \tag{5.11}$$

ensures that each user is active only in a single subband.

Similarly to the scalability constraint defined in (5.5), the number of subcarriers has to be a multiple of 15 kHz:

$$\Delta N_{\Delta f_s}^{u} = 2^{s-1} c_s, \quad s \in \{1,2,3,4\}, c_s \in \mathbb{N}_0. \tag{5.12}$$

The number of subcarriers, if the corresponding user is active, is upper bounded by the total number of subcarriers, $K_{\Delta f_s}^{\max}$, and lower bounded by one:

$$a_{\Delta f_s}^{u} \leq \Delta N_{\Delta f_s}^{u} \leq K_{\Delta f_s}^{\max} a_{\Delta f_s}^{u}. \tag{5.13}$$

Notice that this constraint additionally implies that $\Delta N_{\Delta f}^{u}$ is equal to zero if user u is not active within subband Δf_s.

Users that employ the same numerology are mutually orthogonal. This implies that there is no interference between them, and thus no guard band is needed:

$$N_{\Delta f_s}^{u+1,l} - N_{\Delta f_s}^{u,h} = 0, \tag{5.14}$$

where $u+1$ refers to the user that lies right next to user u[2].

However, when two users with different subcarrier spacings lie next to each other, a guard band is needed since the orthogonality is lost:

$$N_{\Delta f_{s'}}^{u',l} - N_{\Delta f_s}^{u,h} \geq G_{\Delta f_s, \Delta f_{s'}} a_{\Delta f_{s'}}^{u'},\ \Delta f_{s'} > \Delta f_s, u' \neq u. \tag{5.15}$$

[2]Without loss of generality it is assumed that users within the same subband are ordered according to increasing user index.

5.2.2 Resource and Numerology Allocation obtained by applying the LP Algorithm

In order to reduce complexity for larger bandwidth and/or more users, the constraints defined for ILP have to be relaxed. In the following, an integer relaxation to the ILP is applied in order to obtain the LP solution:

$$
\begin{aligned}
&\text{maximize} \quad z \\
&\text{subject to} \quad \sum_{\Delta f_s \in \mathcal{S}} \log_2(1+\bar{\gamma}^u_{\Delta f_s}) \Delta N^u_{\Delta f_s} \geq z, \;\; \Delta f_s \in \mathcal{S}, u \in \mathcal{U}, z \in \mathbb{R} \\
&\qquad N^{u,l}_{\Delta f_s} \leq N^{u,h}_{\Delta f_s}, \;\; \forall u \in \mathcal{U}, \;\; \{N^{u,l}_{\Delta f_s}, N^{u,h}_{\Delta f_s}\} \in \mathbb{R}^+_0 \\
&\qquad \sum_{\Delta f_s \in \mathcal{S}} a^u_{\Delta f_s} = 1, \;\; a_{\Delta f,u} \in [0,1] \\
&\qquad a^u_{\Delta f_s} \leq \Delta N^u_{\Delta f_s} \leq K^{\max}_{\Delta f_s} a^u_{\Delta f_s}, \;\; a_{\Delta f,u} \in [0,1] \\
&\qquad N^{u+1,l}_{\Delta f_s} - N^{u,h}_{\Delta f_s} = 0, \\
&\qquad \Delta N^u_{\Delta f_s} = 2^{s-1} c_s, \;\; s \in \{1,2,3,4\}, c_s \in \mathbb{R}^+_0,
\end{aligned}
\tag{5.16}
$$

Thereby, an integer-relaxation is applied to the following parameters:

$$
N^{u,l}_{\Delta f}, N^{u,h}_{\Delta f}, c_s \in \mathbb{R}^+_0. \tag{5.17}
$$

Additionally, the binary variable $a^u_{\Delta f_s}$ is relaxed so that now it takes any number from the interval $[0,1]$:

$$
a^u_{\Delta f_s} \in [0,1]. \tag{5.18}
$$

Therefore, in general, the solution of (5.16) does not satisfy the constraints of problem (5.8). Specifically, the users might be assigned to multiple subbands and the subband sizes might not be an integer multiple of 15 kHz. Thus, in order to obtain a feasible, though, in general, suboptimal solution to problem (5.8), the solution obtained by solving problem (5.16) has to be transformed into an integer-valued solution. All necessary steps are presented in Algorithm 1 and explained in detail in the following.

1. By solving (5.16), the total number of subcarriers allocated to each user and for each subband ($\Delta N^u_{\Delta f_s}$) is defined. All these values are stacked in matrix $\mathsf{X}^{N_{\text{users}} \times n_{\Delta f_s}}$, where $n_{\Delta f_s}$ is the total number of considered subbands (numerologies) from set $\mathcal{S}$ and N_{users} is the total number of users.

2. Since the elements of X are real values, they have to be first rounded to the nearest integer values.

3. If multiple numerologies are active, it is necessary to assign the guard band between them. If there are subcarriers allocated to any of users within a specific subband, that subband is active. The guard band values are already predefined by setting an upper bound on the allowed INI, i.e., $\sigma^2_{\mathrm{INI}^{\max}}$.

4. Since the optimization does not account for the guard subcarriers, it is necessary to uniformly rescale the total bandwidth by a certain factor, so that the guard band is adequately accounted for.

5. Each allocated subband has to be a multiple of 15 kHz.

6. The final integer-valued outputs representing the total number of subcarriers allocated to each of users within a corresponding subband are stacked in the output matrix $\mathsf{Y}^{N_{\mathrm{users}} \times n_{\Delta f_s}}$.

5.2.3 Resource and Numerology Allocation obtained by applying Heuristic Approaches

In this subsection, two different heuristic approaches for the subcarrier allocation are applied. Both approaches are based on the predefined subcarrier spacing according to:

$$\frac{f_d}{\Delta f_s} \approx \frac{\xi}{T}. \tag{5.19}$$

This equation indicates that the impact of the Doppler shift on the optimal subcarrier spacing is approximately equal to the impact of the RMS delay spread on the optimal symbol duration [62]. By utilizing equation (5.19), it is not always possible to obtain an integer value of a subcarrier spacing Δf_s. Since our work limits the set of possible subcarrier spacings on set $\mathcal{S}$, additionally it is necessary to round the given subcarrier spacing to the nearest possible subcarrier spacing from set $\mathcal{S}$ of all subcarrier spacings.

In the first approach, the optimization problem of (5.8) is run to optimally divide the bandwidth amongst users. This approach is called *heuristic I*. The second approach is *heuristic II*. It simply divides the total bandwidth onto the allocated subbands according to the number of users that are assigned to the subbands. The *heuristic I* approach relies on the integer algorithm explained before, but excludes the third constraint defined in (5.8). The reason for this is that the distribution of users

Algorithm 1: Relaxed solution for equalization of the users' achievable rates

Input : Number of possible subcarrier spacings from the set $\mathcal{S}$, $n_{\Delta f_s}$, total number of subcarriers of 15 kHz spacing $K^{\max}_{\Delta f=15\,\mathrm{kHz}}$, total number of users N_{users}, the output values $\Delta N^u_{\Delta f_s}$ of (5.16) are stacked in matrix $\mathsf{X}^{N_{\mathrm{users}}\times n_{\Delta f_s}}$.

Output : The output matrix containing an integer-valued solution - $\mathsf{Y}^{N_{\mathrm{users}}\times n_{\Delta f_s}}$

Initialization: Set the guard bands to the initial values according to the threshold $\sigma^2_{\mathrm{INI}^{\max}}$. Vector $\mathsf{A}^{1\times n_{\Delta f_s}}$ of all activity variables $a^u_{\Delta f_s}$ is initialized to zero.

1. Round the elements of X ($\Delta N^u_{\Delta f_s}$) to the nearest integer values.

2. In order to satisfy (5.15), the active Δf_s has to be determined. It is done based on the number of assigned subcarriers to each subband and solved by applying a loop shown below. Based on the knowledge about the active Δf_s and $\sigma^2_{\mathrm{INI}^{\max}}$, it is possible to calculate the number of resources assigned to the guard band.

for $s = 1 : n_{\Delta f_s}$ **do**
  if $\mathsf{X}(:, s) \neq 0$ **then**
    $\mathsf{A}(s) = 1$;
  else
    $\mathsf{A}(s) = 0$
  end
end

Based on the active subbands s, the number of guard subcarriers between corresponding subbands is determined. Since we assume four subcarriers spacing from set $\mathcal{S}$, three different guard band sizes can be distinguished: $G_{\Delta f_s,\Delta f_{s'=2s}}$, $G_{\Delta f_s,\Delta f_{s'=4s}}$ and $G_{\Delta f_s,\Delta f_{s'=8s}}$. Accordingly, the activity variables $a_{\Delta f_s,\Delta f_{s'}}$ take on the value one if both subbands s and s' are active ($a^u_{\Delta f_s}$ = 1 and $a^u_{\Delta f'_s}$ = 1); otherwise it is zero.

3. Define the rescaling coefficient α:

$$G_{\mathrm{total}} = G_{\Delta f_s,\Delta f_{s'=2s}} a_{\Delta f_s,\Delta f_{s'=2s}} + G_{\Delta f_s,\Delta f_{s'=4s}} a_{\Delta f_s,\Delta f_{s'=4s}} + G_{\Delta f_s,\Delta f_{s'=8s}} a_{\Delta f_s,\Delta f_{s'=8s}}$$

$$\alpha = \frac{K^{\max}_{\Delta f=15\mathrm{kHz}} - G_{\mathrm{total}}}{K^{\max}_{\Delta f=15\mathrm{kHz}}}$$

4. In order to satisfy (5.12), $\Delta N^u_{\Delta f_s}$ has to be rounded to the nearest integer that is multiple of 2^{s-1}: $\mathsf{Y}(u, \Delta f_s) = [< \alpha \mathsf{X}(u, \Delta f_s) >_{2^{s-1}} == 0]$

is known in advance as well as the subcarrier spacings occupied by the users. In both cases the guard band is considered depending on the chosen subcarrier spacings.

5.2.4 Simulation Results

The simulation results obtained by solving different optimization methods are shown in Fig. 5.4 and Fig. 5.5. Based on the achievable rate $R_{\Delta f_s}$, calculated according to (5.1) and relying on (5.8), the optimal numerology and resource allocation are found in order to equalize the achievable rate of the users. The suboptimal solutions of the LP, the *heuristic I* and the *heuristic II* are compared to the optimal ILP solution. The list of parameters used for obtaining the numerical results is shown in Table 5.2. A bandwidth of 5 MHz, corresponding to a total of 336 subcarriers at 15 kHz, is considered. The TDL-A channel model and RMS delay spread values in the range from 30 ns to 500 ns are used.

Table 5.2: Simulation parameters

parameter	**value**			
Subcarrier spacing [kHz]	15	30	60	120
Number of symbols per subframe	14	28	56	112
CP duration [µs]	4.69	2.34	1.17	0.59
Velocity [km/h]	$\{5, 50, 100, 200, 300\}$			
Doppler spread [Hz]	$\{27.3, 273.3, 546.7, 1093.3, 1640\}$			
Bandwidth [MHz]	5			
Carrier frequency [GHz]	5.9			
Channel model & ξ [ns]	TDL-A $\{30, 50, 100, 200, 300, 400, 500\}$			
Pilot pattern	Diamond			
Number of users	7			

Fig. 5.4 represents the achievable rate normalized to the optimal value achieved by the ILP for three different sets of RMS delay spread values: $\xi_1 \in \{100, 200\}$ ns, $\xi_2 \in \{30, 50, 100, 200\}$ ns and $\xi_3 \in \{30, 50, 100, 200, 300, 400, 500\}$ ns. The RMS delay spread value of each user is randomly selected from these sets and the results are averaged over 50 random realizations. In each realization the minimum achievable rate is considered. The assumed velocity is $v = 100$ km/h for all users. From Fig. 5.4 it can be observed that the suboptimal LP solution is very close to the optimal solution

for all cases. The solution obtained by applying the *heuristic I* algorithm is almost as good as the LP solution. The *heuristic II* approach only performs well when a small set of possible RMS delay spread values are considered. In this case, the users are mostly just assigned to one or two different subbands. For the large set of RMS delay spread values it is necessary to switch among more subcarrier spacings and sharing the bandwith according to the number of users that occupy the same subcarrier spacing results in poor performance. Additionally, the result of an LTE-compliant system is included in order to evaluate the gain of the mixed numerology. In the LTE-compliant system all users are assigned to 15 kHz subcarrier spacing and the bandwidth is equally shared amongst users. Unlike in mixed numerology, in the LTE-compliant system there are no resources wasted on the guard band, since the subcarriers are orthogonal. On the other hand, T_{CP} is large enough to cover a long RMS delay spread of the channel. Hence, the achievable rate of LTE is not far from the optimal solution for all three sets of RMS delay spread, but is still deficient compared to the mixed numerology gain.

In Fig. 5.5, a similar simulation is examined; yet, here the RMS delay spread of 200 ns is constant and the user velocity varies. Three different sets of velocities are assumed: $v_1 \in \{100{,}200\}$ km/h, $v_2 \in \{5{,}50{,}100{,}200\}$ km/h and $v_3 \in \{5{,}50{,}100{,}200{,}300{,}400\}$ km/h. The LP and the heuristic approaches exhibit similar performance as before compared to the ILP. However, the performance of LTE is significantly worse. This is due to a relatively large ICI and a channel estimation error at higher velocities.

Taking into account all these approaches explained before, the algorithms are compared in terms of computational complexity, shown in Fig. 5.6. As expected, the ILP method is the most complex, while the *heuristic II* has the lowest complexity. The optimal heuristic is the second most complex, since it still requires the solution of an integer program, yet with a smaller number of integer variables as compared to ILP. The LP method is considerably less complex than the ILP method, especially for larger number of users. Therefore, it provides a highly favorable tradeoff between computational complexity and performance.

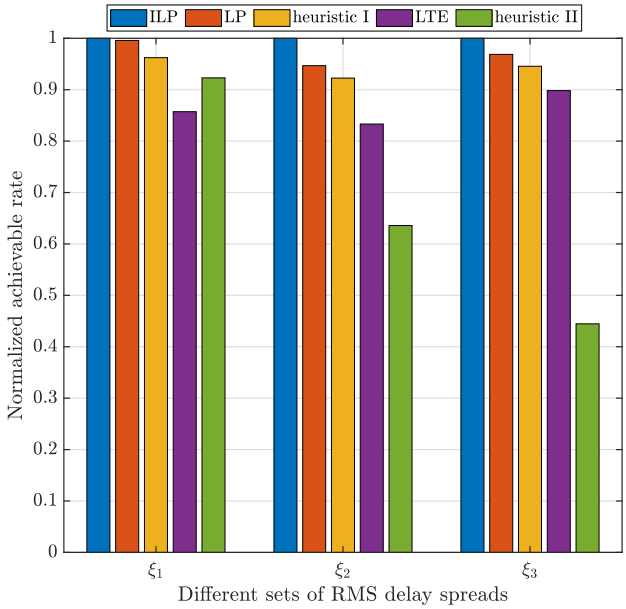


Figure 5.4: Comparison of different methods with constant velocity v = 100 km/h.

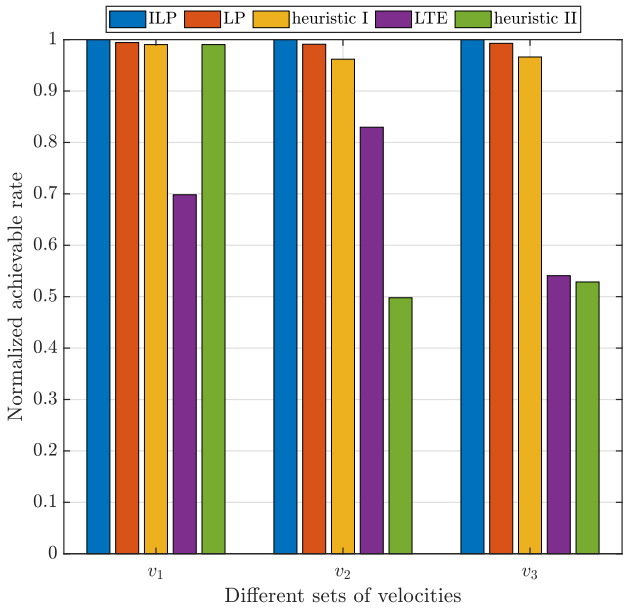


Figure 5.5: Comparison of different methods with constant RMS delay spread ξ = 200 ns.

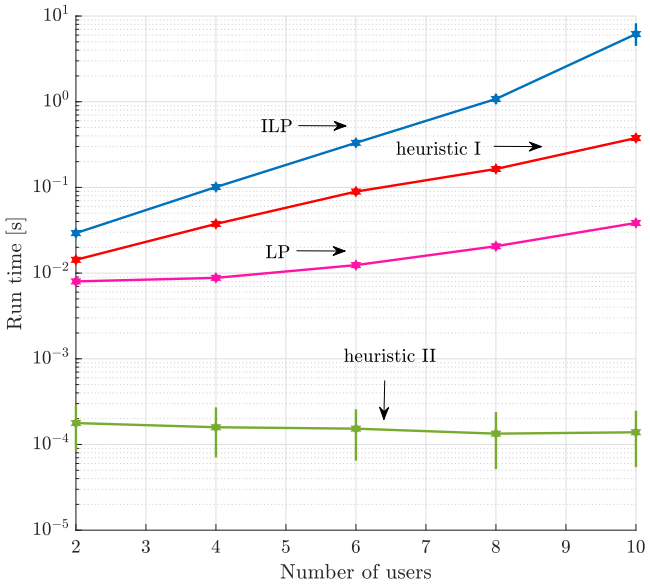


Figure 5.6: Run time for different optimization methods.

Chapter 6

Optimal Resource Allocation based on Mixed Numerology and Mini-slots for Low Latency Communications

3GPP introduces a mixed (multi) numerology approach and a mini-slot approach to enhance the adaptability of the PHY and to support multiple services with different requirements within the same subband [11], [75]. In this section, these two approaches are compared focusing on the achievement of LL communications. The goal is to maximize the achievable rate of best effort users, while maintaining latency requirements of LL users. Exploiting achievable rate performance as one of the fundamental metrics, a comparison of mixed numerology and mini-slot approach in different circumstances is shown. In addition to CP-OFDM, UFMC is employed as a potential beyond 5G technology and an improvement over CP-OFDM is emphasized. The optimization problems for both mixed numerology and the mini-slot approach are initially given by an integer programming solution. In order to reduce computational complexity for large-scale scenarios, the Dantzig-Wolfe decomposition method is applied, showing that it is possible to achieve the optimal solution with significantly reduced complexity.

6.1 Optimization Algorithms for Low Latency Communications

In this part, the constrained optimization for the user resource allocation within appropriate numerology for both, mini-slot and mixed-numerology methods is formulated and explained.

The main criterion on which the optimization is based is the user achievable rate per RE, $\mathrm{R}^{u}_{\Delta f}$, where u denotes a user from set $\mathcal{U}$ of all users. In addition to $\mathrm{R}^{u}_{\Delta f}$, the optimization is constrained with respect to the latency demand τ. In this regard, two groups of users can be distinguished: priority LL users (LLUs) that impose a latency constraint for the transmission of packets of size q, and non LL users (NLLUs) that do not impose a latency constraint. A single LLU is denoted by u_{LL} from set $\mathcal{U}_{\mathrm{LL}}$ of all LLUs and a single NLLU is denoted by u_{NLL} from set $\mathcal{U}_{\mathrm{NLL}}$ of all NLLUs. Set $\mathcal{U}$ comprises

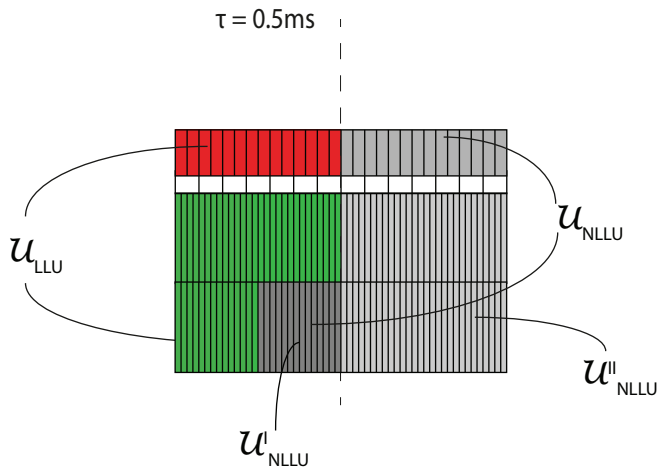
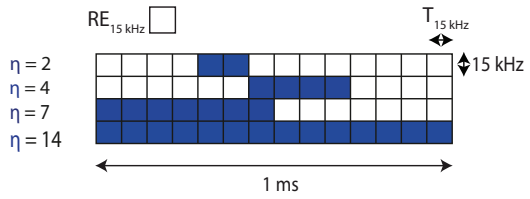


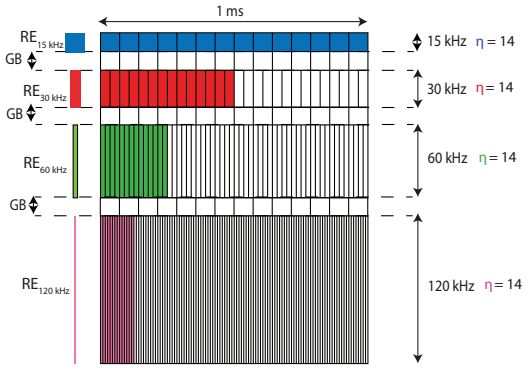
Figure 6.1: An example of the shared resources among LLUs (red and green) and NLLUs (gray), with respect to $\tau = 0.5$ ms.

both groups of users, i.e., $\mathcal{U} = \mathcal{U}_{\text{LL}} \cup \mathcal{U}_{\text{NLL}}$. The total number of LLUs is denoted by $N_{\text{LL}} = |\mathcal{U}_{\text{LL}}|$, and the total number of NLLUs by $N_{\text{NLL}} = |\mathcal{U}_{\text{NLL}}|$. For simplicity, the latency constraint τ takes values equal to the slot duration, $\tau \in \{1, 0.5, 0.25, 0.125\}$ ms that corresponds to the subcarrier spacing from set $\mathcal{S}$, respectively. The same latency constraint τ is assumed for all LLUs in order to simplify the notations; however, this can easily be generalized to user-specific latency constraints. The constraint τ is related to the transmission of packets of size q predefined for LLUs. The same packet size q for all LLUs is considered.

The goal of the optimization is to investigate the optimal numerology and to allocate the optimal number of resources to users with respect to their SINR values and the latency demand τ. An exemplary user allocation is shown in Fig. 6.1. Here, resources shared among LLUs that are represented by red and green colors, and NLLUs that are represented by different shades of gray are considered, with respect to $\tau = 0.5$ ms. This specific case represents the mixed numerology approach with two numerologies (30 kHz and 60 kHz) and a guard band represented by white color. All users that use a given numerology can be pooled in a contiguous subband, since the optimization is based on the average rather than the instantaneous SINR, and therefore it is irrelevant for a user which part of the bandwidth is allocated to it. This simplifies the optimization and it minimizes the guard overhead, as there is no need for multiple guard bands between different numerologies. This can also be applied to the mini-slot approach omitting the guard band. It can be noticed that a more stringent τ reduces the set of possible subcarrier spacings within a mixed numerology approach that can satisfy the latency constraint, i.e., in this specific case it is not possible to select the 15 kHz subcarrier spacing since its slot duration is 1 ms while the latency constraint is 0.5 ms. Similarly, for a mini-slot approach, this implies a reduced set of mini-slot lengths that can be employed to satisfy



(a) Mini-slots can have a length of 2, 4, 7 or 14 symbols.



(b) The slot duration (14 symbols) depends on Δf.

Figure 6.2: An example of mini-slot structure (a) and mixed numerology structure (b).

the latency constraint.

6.1.1 Optimization Algorithms for Low Latency Communications based on Mini-slot

Problem Formulation

The main goal of the optimization is to maximize the minimum achievable rate among NLLUs, ensuring at the same time that the achievable rate of LLUs is large enough to transmit packets of size q within the latency τ. The optimization problem is composed of three parts: the objective function defined in (6.1a), optimization variables defined in (6.1b) and constraints defined in (6.1c). It is given by:

P1:

$$\max \min_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} \quad \sum_{\eta \in \mathcal{W}} \mathrm{R}_{\Delta f}^{u_{\mathrm{NLL}}} N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \eta \tag{6.1a}$$

$$\text{w.r.t.} \quad \Delta f \in \mathcal{S}, N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \in \mathbb{N}_0, N_{\Delta f}^{u_{\mathrm{LL}},\eta} \in \mathbb{N}_0,$$

$$a_{\Delta f}^{u_{\mathrm{NLL}},\eta} \in \{0,1\}, a_{\Delta f}^{u_{\mathrm{LL}},\eta} \in \{0,1\} \tag{6.1b}$$

$$\text{subject to} \quad \sum_{\eta \in \mathcal{W}} \mathrm{R}_{\Delta f}^{u_{\mathrm{NLL}}} N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \eta \geq z, \forall u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}},$$

$$\sum_{\eta \in \mathcal{W}} \mathrm{R}_{\Delta f}^{u_{\mathrm{LL}}} N_{\Delta f}^{u_{\mathrm{LL}},\eta} \eta \geq q, \forall u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}},$$

$$\sum_{\eta \in \mathcal{W}} a_{\Delta f}^{u_{\mathrm{LL}},\eta} = 1,$$

$$a_{\Delta f}^{u_{\mathrm{LL}},\eta} \leq N_{\Delta f}^{u_{\mathrm{LL}},\eta} \leq a_{\Delta f}^{u_{\mathrm{LL}},\eta} N_{\max}^{\eta}, \eta \in \mathcal{W},$$

$$a_{\Delta f}^{u_{\mathrm{NLL}},\eta} \leq N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \leq a_{\Delta f}^{u_{\mathrm{NLL}},\eta} N_{\max}^{\eta}, \eta \in \mathcal{W},$$

$$\sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\mathrm{LL}},\eta} \eta + \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \eta \leq C_{\mathrm{total}},$$

$$\sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\mathrm{LL}},\eta} \eta + \sum_{u_{\mathrm{NLL}}^{I} \in \mathcal{U}_{\mathrm{NLL}}^{I}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\mathrm{NLL}}^{I},\eta} \eta = C_{\mathrm{LL}},$$

$$\sum_{u_{\mathrm{NLL}}^{II} \in \mathcal{U}_{\mathrm{NLL}}^{II}} N_{\Delta f}^{u_{\mathrm{NLL}}^{II},\eta} \eta = C_{\mathrm{total}} - C_{\mathrm{LL}},$$

$$\sum_{u_{\mathrm{NLL}}^{I} \in \mathcal{U}_{\mathrm{NLL}}^{I}} N_{\Delta f}^{u_{\mathrm{NLL}}^{I},\eta} \eta + \sum_{u_{\mathrm{NLL}}^{II} \in \mathcal{U}_{\mathrm{NLL}}^{II}} N_{\Delta f}^{u_{\mathrm{NLL}}^{II},\eta} \eta,$$

$$= \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \eta, \tag{6.1c}$$

with η being the mini-slot weight from set $\mathcal{W}$ = {2, 4, 7, 14} as shown in Fig. 6.2a. The total number of mini-slots with weight η occupied by a single LLU within the latency constraint τ and for subcarrier spacing Δf is denoted by $N_{\Delta f}^{u_{\mathrm{LL}},\eta}$. Similarly, the total number of mini-slots occupied by a single NLLU is denoted by $N_{\Delta f}^{u_{\mathrm{NLL}},\eta}$. Since the minimum time-resource unit is a RE, the number of mini-slots is multiplied with corresponding weight in order to obtain the total number of REs. In addition to η, $N_{\Delta f}^{u_{\mathrm{LL}},\eta}$ and $N_{\Delta f}^{u_{\mathrm{NLL}},\eta}$, the optimal subcarrier spacing Δf from set $\mathcal{S}$ is investigated. In other words, the optimization is solved for each subcarrier spacing separately, and then the most suitable one is selected in terms of the achievable rate. The binary optimization variables that indicate whether a specific NLLU/LLU transmits packets with weight η or not are denoted by $a_{\Delta f}^{u_{\mathrm{NLL}},\eta}$ and

$a_{\Delta f}^{u_{\mathrm{LL}},\eta}$, respectively. The total number of available resources is denoted by C_{total} and the total number of resources predefined for LLUs is denoted by C_{LL}. They are defined as:

$$\begin{aligned} C_{\mathrm{total}} &= K_{\Delta f} M_{\Delta f}^{1ms}, \\ C_{\mathrm{LL}} &= K_{\Delta f} M_{\Delta f}^{\tau}, \end{aligned} \tag{6.2}$$

where $M_{\Delta f}^{\tau}$ is the number of symbols of duration $T_{\Delta f}$ that can fit in latency τ, $M_{\Delta f}^{1ms}$ is the number of symbols that correspond to 1 ms and $K_{\Delta f}$ is the total number of subcarriers as determined by the bandwidth. The total capacity C_{total} therefore represents the total number of resource elements within a subframe. It is subcarrier spacing independent since the symbol duration $T_{\Delta f}$ and subcarrier spacing Δf are inversely proportional, such that this dependency cancels out in the multiplication in (6.2) [57]. In addition to the total capacity that comprises both groups of users, C_{LL} is a limit for LLUs determined by the latency constraint τ, i.e., C_{LL} represents the total number of REs available within the duration τ. Two groups of NLLUs can be distinguished: the first group $\mathcal{U}_{\mathrm{NLL}}^{I}$, which is interleaved with LLUs and hence shares the resources within C_{LL}, and the second group $\mathcal{U}_{\mathrm{NLL}}^{II}$ with the total number of mini-slots per user $N_{\Delta f}^{u_{\mathrm{NLL}}^{II},\eta}$ which occupies the remaining resources. The total number of mini-slots per user from the first group of NLLUs is denoted by $N_{\Delta f}^{u_{\mathrm{NLL}}^{I},\eta}$ and the total number of mini-slots per user from the second group of NLLUs is denoted by $N_{\Delta f}^{u_{\mathrm{NLL}}^{II},\eta}$.

The problem defined in (6) can be reformulated as a pure maximization problem, by introducing an auxiliary variable z, which is a lower bound on the achievable rate of NLLUs. The optimization variables as well as the constraints are already defined in (6.1b) and (6.1c), respectively. These goals are written as:

P2:

$$\begin{aligned} \text{max} \quad & z \\ \text{w.r.t.} \quad & \text{(6.1b)} \\ \text{subject to} \quad & \sum_{\eta \in \mathcal{W}} \mathrm{R}_{\Delta f}^{u_{\mathrm{NLL}}} N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \eta \geq z, \forall u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}} \\ & \text{(6.1c)}. \end{aligned} \tag{6.3}$$

In the following, the constraints of this optimization problem are discussed in more detail.

Latency Constraint

In addition to the channel conditions, in this work a QoS-aware numerology and resource allocation is considered by imposing LL constraints on LLUs. A LL transmission is characterized by a LL parameter τ that does not explicitly show up in the optimization problem defined in (6.3), but it is

a part of the optimization problem through a parameter C_{LL}. As already discussed and shown in Fig. 6.1, C_{LL} predefines the total number of resources that can be assigned to LLUs. Moreover, the optimization specifies a lower bound on the achievable rate of LLUs, which is related to transmission of the packet size q:

$$\sum_{\eta \in \mathcal{W}} \mathrm{R}_{\Delta f}^{u_{\mathrm{LL}}} N_{\Delta f}^{u_{\mathrm{LL}},\eta} \eta \geq q, \forall u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}. \tag{6.4}$$

Therefore, the base station allocates as many resources to LLUs as they need in order to achieve q within the latency constraint τ. By imposing (6.4), the total number of resources that can be assigned to NLLUs is impacted. It means that a smaller packet size that has to be transmitted by LLUs leaves more available resources for NLLUs, thereby affecting their achievable rate, and vice versa.

Activity Constraint

Only blocks with one unique mini slot length can be assigned to each LLU. Therefore, multiplexing of different mini-slots for a single user is not allowed. This constraint facilitates scheduling and simplify the optimization problem by controlling the positions of the resources of LLUs/NLLUs. It is given by means of the binary optimization variable $a_{\Delta f}^{u_{\mathrm{LL}},\eta}$:

$$\sum_{\eta \in \mathcal{W}} a_{\Delta f}^{u_{\mathrm{LL}},\eta} = 1, \;\; a_{\Delta f}^{u_{\mathrm{LL}},\eta} \in \{0, 1\}. \tag{6.5}$$

NLLUs are not restricted to the blocks with specific mini-slot durations. They can select multiple mini-slot lengths in order to maximize the achievable rate. The binary variable of a NLLU is denoted by $a_{\Delta f}^{u_{\mathrm{NLL}},\eta}$.

Activity variables $a_{\Delta f}^{u_{\mathrm{LL}},\eta}$ and $a_{\Delta f}^{u_{\mathrm{NLL}},\eta}$ are functions of the variables $N_{\Delta f}^{u_{\mathrm{LL}},\eta}$ and $N_{\Delta f}^{u_{\mathrm{NLL}},\eta}$, respectively:

$$a_{\Delta f}^{u_{\mathrm{LL}},\eta} \leq N_{\Delta f}^{u_{\mathrm{LL}},\eta} \leq a_{\Delta f}^{u_{\mathrm{LL}},\eta} N_{\max}^{\eta}, \; \eta \in \mathcal{W}, \tag{6.6a}$$

$$a_{\Delta f}^{u_{\mathrm{NLL}},\eta} \leq N_{\Delta f}^{u_{\mathrm{NLL}},\eta} \leq a_{\Delta f}^{u_{\mathrm{NLL}},\eta} N_{\max}^{\eta}, \; \eta \in \mathcal{W}. \tag{6.6b}$$

Hence, by deciding on the number of assigned blocks, the optimizer decides on the activity binary variables, i.e., if there are resources assigned to a specific user, the activity variable takes the value 1; otherwise, the activity variable takes the value 0. The upper bounds of the total number of mini-slots are determined by the maximum number of blocks of weight η that can be fit into the entire resource grid of a subframe, $N_{\max}^{\eta}$.

Knapsack Capacity Constraints

The optimization problem (6.1a) can be identified with a multiscenario max-min *knapsack* problem as a generalization of the standard *knapsack* problem [76]. C_{total} is an upper bound on the total number of resources occupied by both, LLUs and NLLUs:

$$\sum_{u_{\text{LL}} \in \mathcal{U}_{\text{LL}}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\text{LL}},\eta} \eta + \sum_{u_{\text{NLL}} \in \mathcal{U}_{\text{NLL}}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\text{NLL}},\eta} \eta \leq C_{\text{total}}. \tag{6.7}$$

Uniquely defined Resource Constraint

In cases when C_{LL} is not fully utilized by LLUs, the remaining resources are assigned to NLLUs. Two groups of NLLUs can be distinguished - the first group that shares the resources with LLUs and the second group which occupies the remaining resources:

$$\sum_{u_{\text{LL}} \in \mathcal{U}_{\text{LL}}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\text{LL}},\eta} \eta + \sum_{u_{\text{NLL}}^{I} \in \mathcal{U}_{\text{NLL}}^{I}} \sum_{\eta \in \mathcal{W}} N_{\Delta f}^{u_{\text{NLL}}^{I},\eta} \eta = C_{\text{LL}}, \tag{6.8a}$$

$$\sum_{u_{\text{NLL}}^{II} \in \mathcal{U}_{\text{NLL}}^{II}} N_{\Delta f}^{u_{\text{NLL}}^{II},\eta} \eta = C_{\text{total}} - C_{\text{LL}}. \tag{6.8b}$$

The total number of resources assigned to NLLUs has to correspond to the sum of the resources occupied by the first group of NLLUs and resources occupied by the second group of NLLUs:

$$\begin{aligned} \sum_{u_{\text{NLL}}^{I} \in \mathcal{U}_{\text{NLL}}^{I}} N_{\Delta f}^{u_{\text{NLL}}^{I},\eta} \eta + \sum_{u_{\text{NLL}}^{II} \in \mathcal{U}_{\text{NLL}}^{II}} N_{\Delta f}^{u_{\text{NLL}}^{II},\eta} \eta \\ = \sum_{u_{\text{NLL}} \in \mathcal{U}_{\text{NLL}}} N_{\Delta f}^{u_{\text{NLL}},\eta} \eta. \end{aligned} \tag{6.9}$$

A single NLLU can be simultaneously scheduled on resources assigned prior and after the latency constraint, so that one part of user resources belongs to the first group and the remaining resources belong to the second group of NLLUs.

6.1.2 Optimization Algorithms for Low Latency Communications based on Mixed Numerology

Problem formulation

Similarly to the mini-slot approach, the optimization problem for the mixed numerology is composed of three parts: the objective function defined in (6.10a), optimization variables defined in (6.10b) and

optimization constraints defined in (6.10c). It is given by:
P3:

$$\max \min_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} \quad \mathrm{R}_{\Delta f_s}^{u_{\mathrm{NLL}}} N_{\Delta f_s}^{u_{\mathrm{NLL}}} \eta \tag{6.10a}$$

$$\text{w.r.t.} \quad \Delta f_s \in \mathcal{S}, N_{\Delta f_s}^{u_{\mathrm{NLL}}} \in \mathbb{N}_0, N_{\Delta f_s}^{u_{\mathrm{LL}}} \in \mathbb{N}_0,$$

$$N_{\Delta f_s}^{u_{\mathrm{NLL}}^{I}} \in \mathbb{N}_0, N_{\Delta f_s}^{u_{\mathrm{NLL}}^{II}} \in \mathbb{N}_0, K_{\Delta f_s} \in \mathbb{N}_0$$

$$a_{\Delta f_s}^{u_{\mathrm{NLL}}} \in \{0,1\}, a_{\Delta f_s}^{u_{\mathrm{LL}}} \in \{0,1\}, a_{\Delta f_s}^{\mathrm{gl}} \in \{0,1\}, \tag{6.10b}$$

$$\text{subject to} \quad \mathrm{R}_{\Delta f_s}^{u_{\mathrm{LL}}} N_{\Delta f_s}^{u_{\mathrm{LL}}} \eta \geq q, \forall u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}},$$

$$\sum_{\Delta f_s \in \mathcal{S}} a_{\Delta f_s}^{u_{\mathrm{LL}}} = 1,$$

$$\sum_{\Delta f_s \in \mathcal{S}} a_{\Delta f_s}^{u_{\mathrm{NLL}}} = 1,$$

$$a_{\Delta f_s}^{u_{\mathrm{LL}}} \leq N_{\Delta f_s}^{u_{\mathrm{LL}}} \leq a_{\Delta f_s}^{u_{\mathrm{LL}}} N_{\max}^{\eta=14}, \forall \Delta f_s \in \mathcal{S},$$

$$a_{\Delta f_s}^{u_{\mathrm{NLL}}} \leq N_{\Delta f_s}^{u_{\mathrm{NLL}}} \leq a_{\Delta f_s}^{u_{\mathrm{NLL}}} N_{\max}^{\eta=14}, \forall \Delta f_s \in \mathcal{S},$$

$$a_{\Delta f_s}^{\mathrm{gl}} \leq \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} a_{\Delta f_s}^{u_{\mathrm{NLL}}} + \sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}} a_{\Delta f_s}^{u_{\mathrm{LL}}} \leq a_{\Delta f_s}^{\mathrm{gl}} N_{\mathrm{users}},$$

$$\sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}} \sum_{\Delta f_s \in \mathcal{S}} N_{\Delta f_s}^{u_{\mathrm{LL}}} \eta + \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} \sum_{\Delta f_s \in \mathcal{S}} N_{\Delta f_s}^{u_{\mathrm{NLL}}} \eta$$

$$+ \eta G \Big(\sum_{\Delta f_s} a_{\Delta f_s}^{\mathrm{gl}} - 1 \Big) \leq C_{\mathrm{total}},$$

$$\sum_{\Delta f_s} K_{\Delta f_s} 2^{s-1} \leq K_{15\,\mathrm{kHz}} - \sum_{\Delta f_s} a_{\Delta f_s}^{\mathrm{gl}} G + G,$$

$$\sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}^{s}} N_{\Delta f_s}^{u_{\mathrm{LL}}} \eta + \sum_{u_{\mathrm{NLL}}^{I} \in \mathcal{U}_{\mathrm{NLL}}^{I,s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}^{I}} \eta = M_{\Delta f_s}^{\tau} K_{\Delta f_s},$$

$$\sum_{u_{\mathrm{NLL}}^{II} \in \mathcal{U}_{\mathrm{NLL}}^{II,s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}^{II}} \eta = \Big(M_{\Delta f_s}^{1ms} - M_{\Delta f_s}^{\tau} \Big) K_{\Delta f_s},$$

$$\sum_{u_{\mathrm{NLL}}^{I} \in \mathcal{U}_{\mathrm{NLL}}^{I,s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}^{I}} {}^{u_{\mathrm{NLL}}^{I}} + \sum_{u_{\mathrm{NLL}}^{II} \in \mathcal{U}_{\mathrm{NLL}}^{II,s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}^{II}} {}^{u_{\mathrm{NLL}}^{II}}$$

$$= \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}^{s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}}, \tag{6.10c}$$

with $\eta = 14$ being the unique weight for all subcarrier spacings and for all users. Unlike the mini-slot approach, which employs a single subcarrier spacing, the resource allocation within the mixed numerology approach is based on the choice of a suitable subcarrier spacing for each user. Hence, in addition to the optimal number of blocks occupied by LLUs $N_{\Delta f_s}^{u_{\mathrm{LL}}}$ and optimal number of blocks occupied by NLLUs $N_{\Delta f_s}^{u_{\mathrm{NLL}}}$, the optimal numerology for each LLU and NLLU is investigated, i.e., subcarrier spacings Δf_s, where s = {1, 2, 3, 4} corresponds to S = {15 kHz, 30 kHz, 60 kHz 120 kHz}, respectively. In order to obtain the total number of REs, these blocks have to be multiplied with weight η = 14, regardless of subcarrier spacing, as emphasized in Fig. 6.2b. The maximum number of blocks with weight η = 14 is denoted by $N_{\max}^{\eta=14}$. The activity of corresponding numerology is denoted by the binary optimization variable $a_{\Delta f_s}^{\mathrm{gl}}$. The binary optimization variables $a_{\Delta f_s}^{u_{\mathrm{NLL}}}$ and $a_{\Delta f_s}^{u_{\mathrm{LL}}}$ denote the activity of numerology occupied by a NLLU and a LLU, respectively. The total number of resources is divided not only among the total number of blocks occupied by NLLUs/LLUs, but also among the guard band G between numerologies. The total number of subcarriers within a subband bandwidth is denoted by $K_{\Delta f_s}$ and the total number of subcarriers of the 15 kHz subcarrier spacing within the available bandwidth is denoted by $K_{15\,\mathrm{kHz}}$. Similarly to the mini-slot approach, two groups of users can be distinguished: the first group $\mathcal{U}_{\mathrm{NLL}}^{I,s}$ which shares the resources with LLUs within the total capacity of LLUs C_{LL}, and the second group $\mathcal{U}_{\mathrm{NLL}}^{II,s}$ which occupies the remaining resources. The total number of blocks per user from the first group of NLLUs is denoted by $N_{\Delta f_s}^{u_{\mathrm{NLL}}^{I}}$ and the total number of blocks per user from the second group of NLLUs is denoted by $N_{\Delta f_s}^{u_{\mathrm{NLL}}^{II}}$.

Following the idea of (6.3), the problem given in (15) can be reformulated as:

P4:

$$\begin{aligned} &\max && z \\ &\text{w.r.t.} && (6.10b) \\ &\text{subject to} && \sum_{\Delta f_s \in \mathcal{S}} \mathrm{R}_{\Delta f_s}^{u_{\mathrm{NLL}}} N_{\Delta f_s}^{u_{\mathrm{NLL}}} \eta \geq z, \forall u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}} \\ & && (6.10c). \end{aligned} \tag{6.11}$$

Latency Constraint

Similarly to 6.1.1, the latency constraint is imposed within the mixed numerology approach:

$$\mathrm{R}_{\Delta f_s}^{u_{\mathrm{LL}}} N_{\Delta f_s}^{u_{\mathrm{LL}}} \eta \geq q, \forall u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}. \tag{6.12}$$

It is important to mention that in this case a LL parameter τ impacts the number of numerologies that the base station can assign to users. The smaller this parameter, the narrower the set of possible numerologies.

Activity Constraints

Within the mixed numerology approach, users can choose between all possible subcarrier spacings from set $\mathcal{S}$, but each LLU/NLLU can select only one optimal subcarrier spacing. In such a way, all resources of a specific user are consolidated within a single subcarrier spacing. The corresponding constraints are given by means of the binary optimization variables $a_{\Delta f_s}^{u_{\mathrm{LL}}}$ for LLUs and $a_{\Delta f_s}^{u_{\mathrm{NLL}}}$ for NLLUs:

$$\sum_{\Delta f_s \in \mathcal{S}} a_{\Delta f_s}^{u_{\mathrm{LL}}} = 1, \;\; a_{\Delta f_s}^{u_{\mathrm{LL}}} \in \{0,1\} \tag{6.13a}$$

$$\sum_{\Delta f_s \in \mathcal{S}} a_{\Delta f_s}^{u_{\mathrm{NLL}}} = 1, \;\; a_{\Delta f_s}^{u_{\mathrm{NLL}}} \in \{0,1\}. \tag{6.13b}$$

These binary variables show the activity status of corresponding LLU/NLLU within subcarrier spacing Δf_s. If the user is active, this variable takes on the value one, otherwise it is zero.

Similarly to the mini-slot approach, these activity variables are determined by the number of assigned blocks to LLUs/NLLUs:

$$a_{\Delta f_s}^{u_{\mathrm{LL}}} \leq N_{\Delta f_s}^{u_{\mathrm{LL}}} \leq a_{\Delta f_s}^{u_{\mathrm{LL}}} N_{\max}^{\eta=14}, \; \forall \Delta f_s \in \mathcal{S}, \tag{6.14a}$$

$$a_{\Delta f_s}^{u_{\mathrm{NLL}}} \leq N_{\Delta f_s}^{u_{\mathrm{NLL}}} \leq a_{\Delta f_s}^{u_{\mathrm{NLL}}} N_{\max}^{\eta=14}, \; \forall \Delta f_s \in \mathcal{S}. \tag{6.14b}$$

In addition to the user activity variable, the binary variable $a_{\Delta f_s}^{\mathrm{gl}}$ that refers to the activity status of subcarrier spacing Δf_s is defined. More precisely, this global binary variable indicates whether a specific subcarrier spacing is occupied by any of the users (LLUs and NLLUs) or not. Therefore, it is not a characteristic of a single LLU/NLLU, but rather a characteristic of a subband that might contain multiple LLUs and/or NLLUs. Thus, the global indicator variable $a_{\Delta f_s}^{\mathrm{gl}}$ of a given subcarrier spacing Δf_s has to be one if any user occupies this subcarrier spacing and zero otherwise:

$$a_{\Delta f_s}^{\mathrm{gl}} \leq \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} a_{\Delta f_s}^{u_{\mathrm{NLL}}} + \sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}} a_{\Delta f_s}^{u_{\mathrm{LL}}} \leq a_{\Delta f_s}^{\mathrm{gl}} N_{\mathrm{users}}. \tag{6.15}$$

Knapsack Capacity Constraints

Below, an upper bound on the available resources for both groups of users within a given bandwidth in the frequency domain and 1 ms of time span is defined.

Having a constant weight η for different numerologies implies that the same number REs can be placed within a total capacity C_{total}. This allows to define C_{total} and C_{LL} in the same way as in

(6.2). The total capacity does not only include the resources occupied by LLUs and the resources occupied by NLLUs, but also the guard band between different numerologies, G:

$$\sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}} \sum_{\Delta f_s \in \mathcal{S}} N_{\Delta f_s}^{u_{\mathrm{LL}}} \eta + \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}} \sum_{\Delta f_s \in \mathcal{S}} N_{\Delta f_s}^{u_{\mathrm{NLL}}} \eta + \eta G \Big(\sum_{\Delta f_s} a_{\Delta f_s}^{\mathrm{gl}} - 1 \Big) \leq C_{\mathrm{total}}. \tag{6.16}$$

Based on the threshold that is set as an upper bound on allowed INI, as explained in Chapter 4, the size of the guard band is determined with respect to the 15 kHz subcarrier spacing in the frequency domain.

Uniquely defined Resource Constraint

This constraint checks that the total assigned bandwidth does not exceed the total available bandwidth. To achieve this, the following constraint is introduced:

$$\sum_{\Delta f_s} K_{\Delta f_s} 2^{s-1} \leq K_{15\,\mathrm{kHz}} - \sum_{\Delta f_s} a_{\Delta f_s}^{\mathrm{gl}} G + G, \tag{6.17}$$

where $K_{\Delta f_s}$ is the total number of subcarriers within subband s, $K_{15\,\mathrm{kHz}}$ is the total number of subcarriers of the 15 kHz subcarrier spacing within the available bandwidth and $a_{\Delta f_s}^{\mathrm{gl}}$ is the binary variable that shows the activity of corresponding numerology Δf_s. In this case, $K_{\Delta f_s}$ is an optimization variable. Since the frequency domain allocation is performed with respect to 15 kHz, the number of subcarriers $K_{\Delta f_s}$ has to be scaled by a factor 2^{s-1}.

It can happen that not all available resources within C_{LL} are occupied by LLUs and therefore, the remaining resources are assigned to NLLUs. Again, two groups of NLLUs can be distinguished: the first group $\mathcal{U}_{\mathrm{NLL}}^{I}$ which shares the resources with LLUs within C_{LL}, and the second group $\mathcal{U}_{\mathrm{NLL}}^{II}$ which occupies the remaining resources. The following equations are given with respect to the subset of LLUs/NLLUs that are assigned to a specific subband s. The set of all LLUs which correspond to the same subband s is denoted by $\mathcal{U}_{\mathrm{LL}}^{s}$. Similarly, sets of NLLUs from the first and second group and the same subband are denoted by $\mathcal{U}_{\mathrm{NLL}}^{I,s}$ and $\mathcal{U}_{\mathrm{NLL}}^{II,s}$, respectively. For the resources that lie below τ, the following constraint is defined:

$$\sum_{u_{\mathrm{LL}} \in \mathcal{U}_{\mathrm{LL}}^{s}} N_{\Delta f_s}^{u_{\mathrm{LL}}} \eta + \sum_{u_{\mathrm{NLL}}^{I} \in \mathcal{U}_{\mathrm{NLL}}^{I,s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}^{I}} \eta = M_{\Delta f_s}^{\tau} K_{\Delta f_s}, \ \forall \Delta f_s \in \mathcal{S}. \tag{6.18}$$

For the remaining resources a similar constraint is defined:

$$\sum_{u_{\mathrm{NLL}}^{II} \in \mathcal{U}_{\mathrm{NLL}}^{II,s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}^{II}} \eta = \left(M_{\Delta f_s}^{1ms} - M_{\Delta f_s}^{\tau} \right) K_{\Delta f_s}. \tag{6.19}$$

Additionally, the sum of resources from the first group of NLLUs and resources from the second group of NLLUs at appropriate subcarrier spacing has to correspond to the total number of resources of NLLUs:

$$\begin{aligned} \sum_{u_{\mathrm{NLL}}^{I} \in \mathcal{U}_{\mathrm{NLL}}^{I,s}} N_{\Delta f_s^{u_{\mathrm{NLL}}^{I}}}^{u_{\mathrm{NLL}}^{I}} + \sum_{u_{\mathrm{NLL}}^{II} \in \mathcal{U}_{\mathrm{NLL}}^{II,s}} N_{\Delta f_s^{u_{\mathrm{NLL}}^{II}}}^{u_{\mathrm{NLL}}^{II}} \\ = \sum_{u_{\mathrm{NLL}} \in \mathcal{U}_{\mathrm{NLL}}^{s}} N_{\Delta f_s}^{u_{\mathrm{NLL}}}. \end{aligned} \tag{6.20}$$

6.1.3 Performance Comparison

In this section, the performance of the optimization methods is exhibited by numerical simulations. A multi-user scenario of N_{users} = 20 - N_{LLU} = 10 and N_{NLLU} = 10 is assumed. Each user experiences different channel conditions characterized by the Doppler and delay spread of the channel. The simulation parameters are summarized in Table 6.1.

Table 6.1: Simulation parameters

s	$s=1$	$s=2$	$s=3$	$s=4$
Subcarrier spacing [kHz]	15	30	60	120
Number of subcarriers	576	288	144	72
Number of resource elements	8064			
Bandwidth [MHz]	8.64			
RMS delay spread [ns]	TDL-A $\{80, 180, 280, 380, 480\}$			
Velocity [km/h]	$\{5, 50, 100, 200, 300\}$			
Doppler spread [Hz]	$\{27.3, 273.3, 546.7, 1093.3, 1640\}$			
Number of users	10 LLUs + 10 NLLUs			
Number of realization	20			
τ [ms]	$1, 0.25$			
Waveform	OFDM, UFMC			
Carrier frequency [GHz]	5.9			

This section considers the results for three different scenarios with respect to the channel conditions employing either OFDM or UFMC. Therefore, the waveform employed by a base station is unique

for all numerologies and it does not vary between them. A TDL-A channel model is assumed for all scenarios, [66]. The number of subcarriers, and therefore the bandwidth, was chosen such that it provides meaningful results in terms of the achievable rate on the one hand and that the solution is obtained with reasonable amount of computing effort on the other hand [77]. The total number of resource elements is calculated as a multiplication of the total number of available subcarriers of corresponding subcarrier spacing and the total number of symbols of corresponding duration within 1 ms, according to (6.2) and Table 6.1. The number of users was selected according to the bandwidth size and packet size q, so that all users can be supported within the available bandwidth and the maximum considered packet size.

In the following, the results imposing different latency constraints τ are compared. Specifically, latency values τ that are equal to the slot durations of 15 kHz and 60 kHz, i.e., 1 ms and 0.25 ms, respectively are selected. The results are provided with respect to two metrics: percentage of served LLUs (β) and sum achievable rate of NLLUs. The LLUs occupy as many resources as they need to convey their packet size q within the latency constraint τ. The remaining resources are assigned to NLLUs. Correspondingly, the value of the packet size q has an impact on the sum achievable rate of NLLUs. Larger q implies more resources for LLUs and hence less resources for NLLUs. However, if q is too large such that not all LLUs can be satisfied within capacity C_{LL}, then some of the LLUs will be discarded by the scheduler; these users are therefore in outage. In order to calculate the percentage of served LLUs, a metric β is defined as:

$$\beta[\%] = \frac{\hat{N}_{\mathrm{LLU}}}{N_{\mathrm{LLU}}} 100, \tag{6.21}$$

with $\hat{N}_{\mathrm{LLU}}$ being the number of satisfied LLUs. In case none of the LLUs can be satisfied, the optimization is infeasible. For the mini-slot approach the users with the lowest SINR are firstly discarded, since they require most of the resources. LLUs are repeatedly discarded, until a feasible solution of (6.3) is obtained. For the mixed numerology approach, users are discarded randomly, since different subcarrier spacings can be assigned to different users and it is therefore not possible to uniquely identify users with lowest SINRs (the SINR depends on the subcarrier spacing).

The first scenario (Scenario I) assumes a fixed RMS delay spread of 200 ns for each user and various Doppler spreads characterized by the velocities from set $\{5, 50, 100, 200, 300\}$ km/h. A velocity from this set is randomly assigned to each user and the optimization results are averaged over 20 random realizations. Fig. 6.3 shows the sum achievable rate of NLLUs as a function of the packet size q of LLUs for τ = 1 ms. In this case, due to the relaxed latency constraint, all LLUs are satisfied for all q and β = 100 %. It can be observed that the different subcarrier spacings achieve very different performance. The best performance in Scenario I is achieved by 30 kHz, which provides on average the highest SINR for the employed set of channel characteristics. Subcarrier spacings of 15 kHz and 120 kHz perform worst in Scenario I: 15 kHz exhibits strong ICI at larger velocities, whereas 120 kHz exhibits ISI for 200 ns of RMS delay spread. For the mini-slot approach, only CP-OFDM

is simulated, since with mini-slots there is no INI and therefore CP-OFDM and UFMC achieve the same performance. However, in the case of mixed numerology, two different subcarrier spacings that achieve the best SINR values (30 kHz and 60 kHz) are assigned to users. In this case, INI exists and some resources are assigned to the guard band. From Fig. 6.3 it can be observed that the mixed numerology method achieves similar performance as the mini-slot approach for 30 kHz, irrespective of the overhead caused by the guard bands and the smaller SINR values due to INI. Employing UFMC in the mixed numerology case provides only a minor gain. This gain is due to smaller guard bands, which are reduced from 12 subcarriers for OFDM to 8 subcarriers for UFMC. The reason for very similar guard overhead is that the interference of both systems close to the border between different numerologies is very similar as shown in Fig. 4.5.

Fig. 6.4 shows the percentage of served LLUs for $\tau = 0.25$ ms under Scenario I. It can be observed that in this case not all LLUs can be served within the given LL capacity C_{LL} with growing packet size q; therefore, the scheduler has to discard some of these users, decreasing β. In addition, β is also impacted by the SINRs of the users; lower SINR implies smaller β and higher outage probability. Interestingly, the best performance for LLUs is not achieved by 120 kHz subcarrier spacing, as one could expect, due to the number of slots that can be accommodated within this latency constraint. The reason is the low SINR values due to the imposed channel conditions. Hence, it can be noticed that this subcarrier spacing experiences the lowest level of served LLUs.

In Fig. 6.5, the sum achievable rate for $\tau = 0.25$ ms under Scenario I is shown. The order of the curves matches the one in Fig. 6.3. Unlike in Fig. 6.3 where curves have a linear behaviour, here it can be noticed that after a certain packet size, the sum achievable rate saturates. The reason for this is that from a certain q on, all resources that are predefined for LLUs, C_{LL}, are fully occupied and therefore the remaining resources assigned to NLLUs stay constant with small variations. Additionally, Fig. 6.5 shows the results obtained applying mixed numerology, employing both OFDM and UFMC. The main difference between the mini-slot approach and the mixed numerology approach is that former employs one numerology with different slot durations (mini-slots) and latter employs multiple numerologies with their regular slot durations according to Table 2.1. Therefore, by considering τ = 0.25 ms, it is not possible to employ 15 kHz and 30 kHz numerologies whose slot durations are 1 ms and 0.5 ms, respectively. Due to the narrow choice of subcarrier spacings (60 kHz and 120 kHz) that achieve τ = 0.25 ms and poor performance of 120 kHz due to the considered RMS delay spread, the optimization provides the resource allocation based only on 60 kHz subcarrier spacing. In such a case, when only a single subcarrier spacing is employed, UFMC provides the same result as OFDM. The reason for that is a lack of the guard band. Therefore, the mixed numerology case of both OFDM and UFMC and the mini-slot case with 60 kHz subcarrier spacing overlap.

The second scenario (Scenario II) assumes a fixed Doppler spread of 1.093 kHz, (200 km/h velocity) and Jakes' Doppler spectrum, and various RMS delay spreads from set $\{80, 180, 280, 380, 480\}$ ns. The sum achievable rate of NLLUs under Scenario II with τ = 1 ms and τ = 0.25 ms, as well as, the percentage of served LLUs are shown in Fig. 6.6, Fig. 6.8 and Fig. 6.7, respectively. In addition to

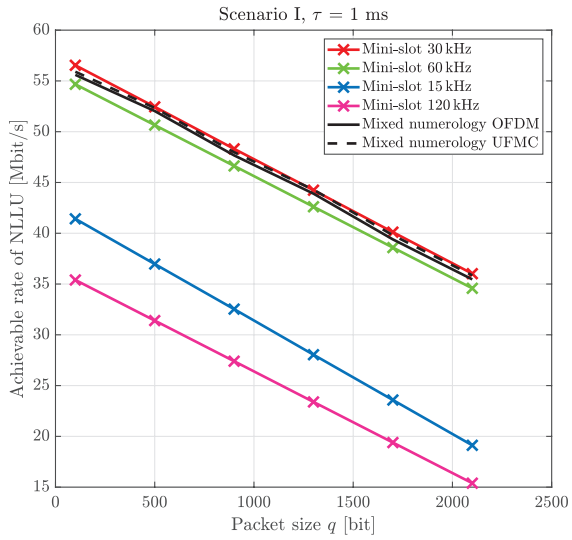


Figure 6.3: Sum achievable rate of NLLUs for Scenario I and $\tau = 1$ ms.

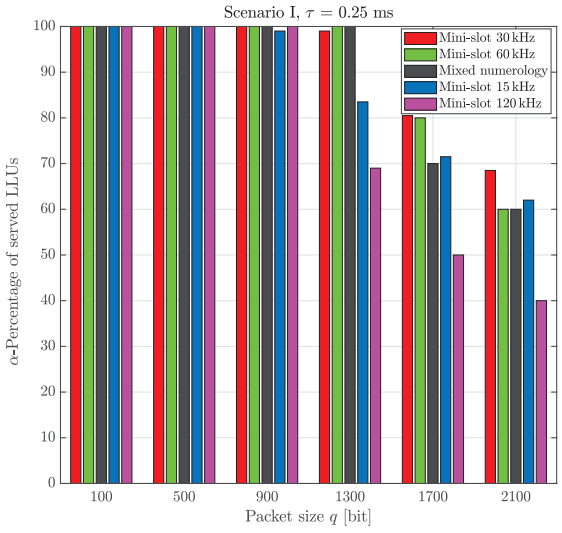


Figure 6.4: Percentage of served LLUs for Scenario I and $\tau = 0.25$ ms.

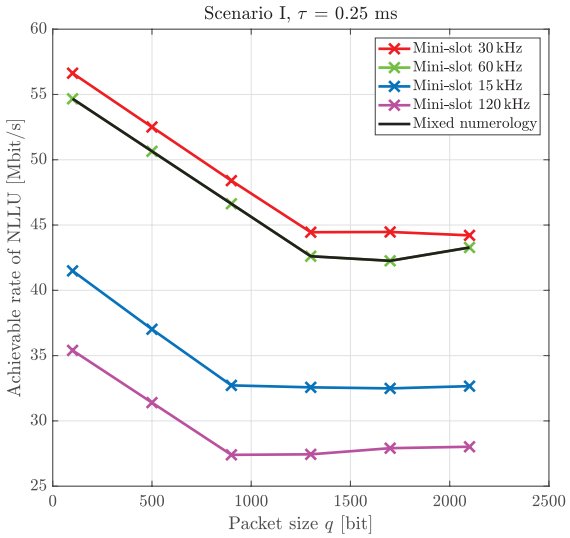


Figure 6.5: Sum achievable rate of NLLUs for Scenario I and $\tau = 0.25$ ms.

the lower sum achievable rate of NLLUs compared to the same τ under Scenario I, the saturation can be noticed already with small packet sizes. This indicates that Scenario II is less suitable than Scenario I for the employed numerology. However, the order of the curves remains unchanged with an optimal $\Delta f = 30$ kHz in both cases.

The third scenario (Scenario III) assumes opposing channel conditions: one group of users experience large RMS delay spreads from set $\{380, 480\}$ ns and small Doppler spread of 273.337 Hz (50 km/h velocity), whereas the other group of users experience large Doppler spread of 1.093 kHz (200 km/h velocity) and RMS delay spread from set $\{80, 180\}$ ns. Fig. 6.9 shows the sum achievable rate of NLLUs under Scenario III with the latency constraint $\tau = 1$ ms. In this case, mixed numerology shows better adjustment to different channel conditions and hence achieves better performance than a single numerology. The group of users that experience small Doppler spread and large RMS delay spread suffer from large ISI. Therefore, large subcarrier spacings are not suitable for this group of users. On the other hand, small subcarrier spacings are not convenient for large ICI caused by large Doppler spread for the other group of users. Additionally, it can be observed that 15 kHz subcarrier spacing exhibits larger achievable rate compared to Scenario I and Scenario II due to the large RMS delay spreads. Similarly to these scenarios, 30 kHz subcarrier spacing manifests the best characteristics among numerologies.

Summarizing this section, the main insights into the results are provided.

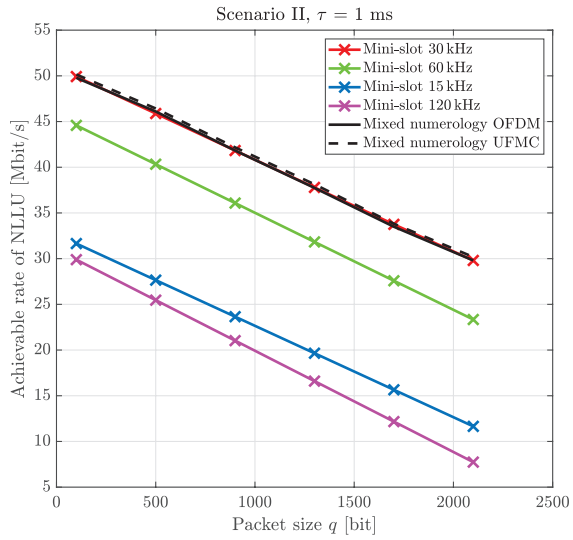


Figure 6.6: Sum achievable rate of NLLUs for Scenario II and τ = 1 ms.

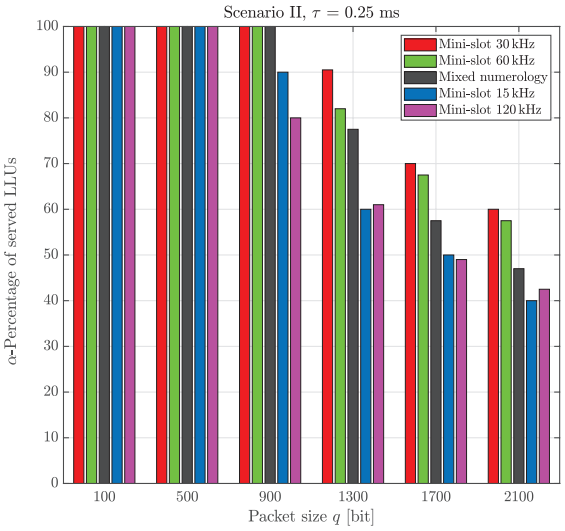


Figure 6.7: Percentage of served LLUs for Scenario II and τ = 0.25 ms.

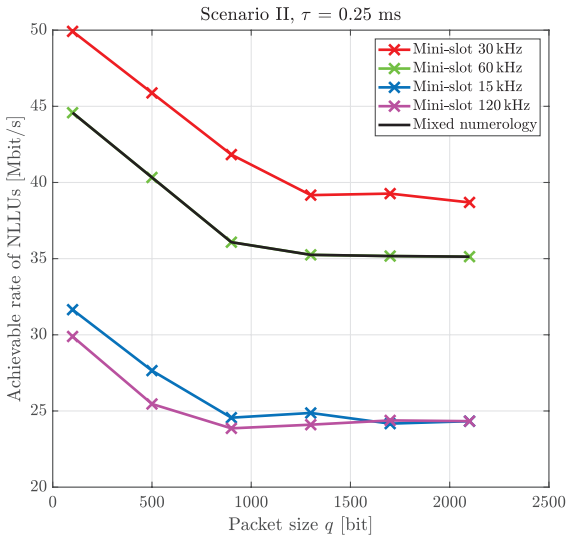


Figure 6.8: Sum achievable rate of NLLUs for Scenario II and τ = 0.25 ms.

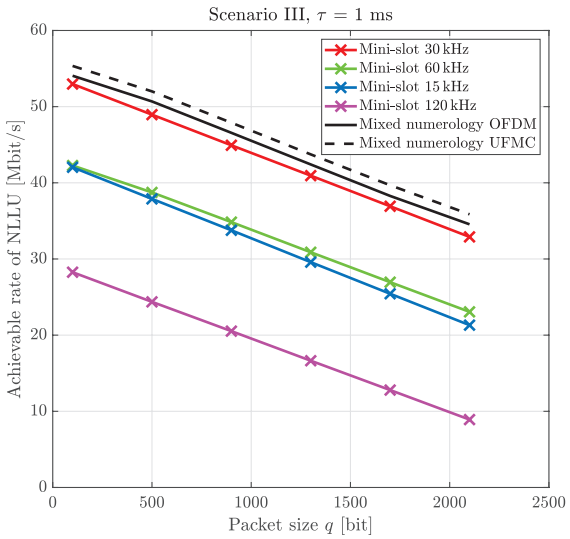


Figure 6.9: Sum achievable rate of NLLUs for Scenario III and τ = 1 ms.

- In cases of a flexible latency constraint two different cases can be distinguished:
  - Both methods perform similarly when non-conflicting channel conditions are observed.
  - Mixed numerology outperforms the mini-slot approach when quite opposing channel conditions are observed.
- In cases of a stringent latency constraint, the mini-slot approach outperforms mixed numerology.

6.2 Dantzig-Wolfe Decomposition for Computational Complexity Reduction

So far, it has been shown how the previously explained optimization problems are defined. These optimization problems can be solved using any ILP solver. However, ILP is NP-hard. Additionally to this optimizer, in this work the *Dantzig-Wolfe* decomposition method followed by a *column generation* algorithm is applied. By enabling a parallel execution of the set of mutually independent linear programming problems, the Dantzig-Wolfe decomposition provides the optimal solution with reduced computational complexity of the algorithm for large-scale scenarios compared to the two previously explained approaches [78].

6.2.1 Overview of the Dantzig-Wolfe Decomposition

The *Dantzig-Wolfe* algorithm decomposes the original problem into a number of pricing or subproblems and a master problem that couples and coordinates the subproblems [79]. The idea of decomposition into subproblems is to obtain a solver-friendly form. This is achieved by carrying out the optimization on a smaller set of variables compared to the ILP techniques, where the size of the original problem is larger, implying a more computationally expensive execution. In order to employ the *Dantzig-Wolfe* algorithm, the constraints of the optimization problem have to be decomposable in a primal block angular structure. In considered case, this can be achieved by partitioning the constraints into those relating to NLLUs and those relating to LLUs, as shown in Fig. 6.10. These subblocks are formed from the constraints that depend only on the optimization variables either connected to NLLUs or to LLUs. On top of these blocks, linking (master) constraints that couple the optimization variables of both NLLUs and LLUs are accommodated. To clarify this, an example of

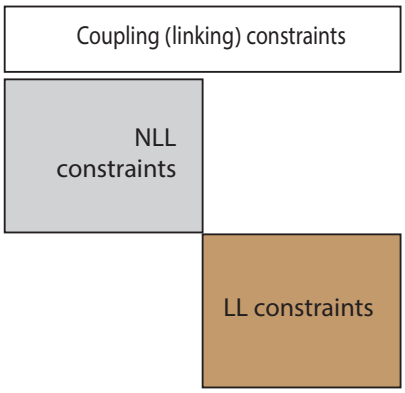


Figure 6.10: Primal block angular structure of the constraint matrix.

the optimization method for the mini-slot approach defined in Sec. 6.1.1 is considered. A similar approach can be applied in the mixed numerology case. The constraints are split as follows:

1. The group of NLL constraints involves the first constraint of (6.3) that is relevant for NLLUs, as well as, (6.6b), (6.8b) and (6.9).

2. The group of LL constraints involves (6.4), (6.5) and (6.6a).

3. The linking constraints are defined in (6.7) and (6.8a).

According to the *Minkowski's representation theorem*, each of these subblocks can be observed as a polyhedron with its extreme points and rays. Using this fact, a master problem can be reformulated in such a way that a feasible region given by a linear constraint representation is replaced by a convex combination of its extreme points and rays. Since a bounded polyhedron is observed, there is no need to consider the extreme rays, but only the extreme points. The number of master associated variables that correspond to the extreme points of the subproblem can be exponentially large; therefore instead of the full-size master problem, only a restricted master problem is solved. It implies that the optimization executes only on the set of the extreme points using the *column generation* algorithm [78, 79]. Based on the optimal dual solution of the restricted master problem suitable objective functions of the subproblems are defined. By solving this, an objective value on each extreme point is obtained and that is considered to be a reduced cost of the restricted master problem. The profitable reduced costs are added to the current basis. The profitable reduced cost implies the value by which an objective function coefficient has to be improved. Since the maximization problem is solved, the maximum reduced cost is considered. If that value is positive then the associated master problem value enters the basis. This process of finding new extreme points (columns) iterates until there are no variables that enter the basis, signifying that an optimal solution has been found. A step-wise explanation of this algorithm is provided in Appendix 9.

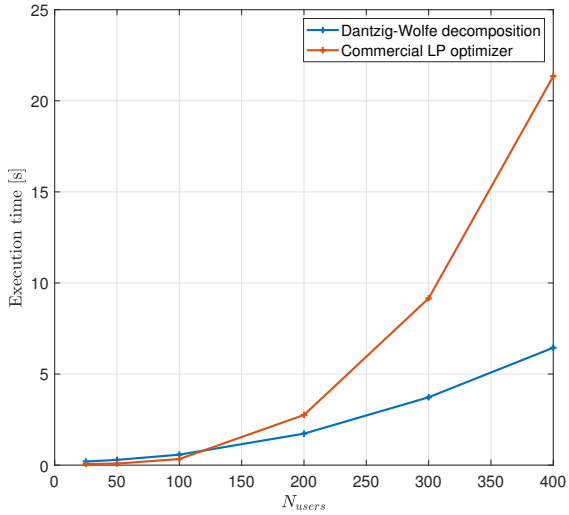


Figure 6.11: Comparison of the commercial LP optimizer vs. DW decomposition.

In order to reduce computational complexity of the optimization for large-scale scenarios and obtain the optimal solution at the same time, in this work the importance of employing more sophisticated optimization methods such as the *Dantzing-Wolfe* decomposition is investigated. In Fig. 6.11, the commercial LP[1] optimizer embedded in the Matlab optimization toolbox and the *Dantzing-Wolfe* decomposition followed by the *column generation* algorithm are compared. The computational complexity of both solvers is the same and it is $\mathcal{O}(m^{2,38})$, where m is the input size of the optimization problem [80]. The gain visible in Fig. 6.11 arises due to the parallel execution of two independent sub-problems whose dimensions are smaller than m. The dimensions of the subproblems depend on the number of LLUs/NLLUs. If a dimension m_1 of the first subproblem that is related to NLLUs and $m_1 < m$, then its complexity is given by $\mathcal{O}(m_1^{2,38})$ and similarly, for the second subproblem it is $\mathcal{O}(m_2^{2,38})$, where $m_2 < m$. It holds true that $\mathcal{O}(m_1^{2,38} + m_2^{2,38}) < \mathcal{O}((m_1 + m_2)^{2,38}) = \mathcal{O}(m^{2,38})$. Of course, the computational complexity within the Dantzig-Wolfe algorithm depends on the number of iterations that have to be executed in order to find an optimal solution. However, only one iteration is sufficient to achieve this in considered case. The results are given with respect to the execution time of the optimization versus the number of users. It can be observed that the *Dantzing-Wolfe* decomposition provides significantly better results for a large number of users due to the parallel execution on the set of optimization variables, rather than on the full-size problem. The commercial LP solvers are sophisticated enough for small number of users.

[1] Instead of an ILP solver that exhibits very high computational complexity, in this example an LP solver is employed, providing an approximate solution with smaller complexity.

Chapter 7

Conclusion

The final chapter of the thesis emphasizes the main contributions and draws the conclusions based on the presented work. The possibilities for improving the existing algorithms and their extension for more complex scenarios are proposed for future research.

7.1 Summary of Contributions

A highly flexible framework in terms of scalable numerology and waveforms has opened a wide range of possible topics within the 5G community. Since 5G NR targets to multiplex various applications with different quality of service requirements within the same band, 3GPP introduces a mixed (multi) numerology approach and a mini-slot approach to enhance the adaptability of the PHY. The thesis focuses on resource allocation and optimal choice of numerology in a multi-user mixed numerology and mini-slot scenario, which is primarily based on the SINR value of a specific user modeled by PHY parameters, as well as on its QoS requirements. In addition to CP-OFDM, which is chosen as the waveform for 5G NR in downlink, UFMC is also considered for beyond 5G technologies and the impact of this waveform on the performance is investigated in the thesis.

Chapter 2 encompasses different research on waveforms over the years, providing detailed insight into the main characteristics of the waveforms used in the thesis, namely CP-OFDM and UFMC, and mutually comparing them. Furthermore, it examines the definition of scalable numerology proposed by 3GPP, emphasizing its main qualities.

Upon entering a new topic to explore, the first goal was to examine the importance of using flexible numerology as well as scenarios in which its gain is visible compared to the previous generations. The main criterion based on which the performance is investigated is the user SINR, since it adequately describes the changes that occur in different scenarios, i.e., channel conditions that each of the users experiences. The SINR closed-form expression accounts for the channel imperfections, such as ISI and ICI, as well as the channel estimation error. Based on the results presented in Chapter 3, it can be concluded that the optimal numerology is determined by the interplay between the aforementioned

parameters. With the goal of maximizing the throughput performance, the optimal pilot patterns are also selected in individual circumstances. As it is visible from the results in Chapter 3, smaller subcarrier spacings are suitable for environments with higher RMS delay spread and small Doppler shifts, for instance urban environment. Nonetheless, for drive-by scenarios where the Doppler effect is pronounced and for lower RMS delay spread values, larger subcarrier spacings are favorable. This is especially convenient for environments such as rural, suburban and highways. In circumstances where it is not entirely clear which optimal numerology should be used, as well as in conditions where a large number of different users is present, the use of mixed numerology becomes even more substantial.

The introduction of mixed numerology implies the destruction of orthogonality between neighbouring users, consequently worsening the performance. In Chapter 4, closed-form expressions of INI are derived for both OFDM and UFMC and this parameter is also included in the SINR expression. In addition to this, the chapter presents how the guard band is defined and how its choice affects INI. The conclusion drawn is that the increase in the number of guard subcarriers reduces INI power. On the other hand, although a large number of guard subcarriers solves the INI problem, it reduces the frequency efficiency of the system due to the increased overhead. Therefore, it is very important to find the appropriate number of guard subcarriers between numerologies, so that a tradeoff between INI and spectral efficiency can be achieved. Additionally, the impact of different filter lengths in the UFMC case is discussed. The results show that INI is reduced if longer filters are applied. However, intrinsic interference is also created when the CP length is insufficient to compensate for both RMS delay spread and the filter tail, which does not serve to improve the performance, quite the contrary.

The most significant contributions of the thesis are presented through two novel optimization algorithms in Chapter 5. The initial step is to examine an algorithm for the optimal resource allocation in a mixed numerology scenario. This method shows how the available bandwidth is shared among users who employ different numerologies under the same channel conditions, in order to equalize their achievable rates. The number of resources assigned to different users depends on the channel conditions, i.e., their SINR values. If the numerology used by a particular user is not fully suitable to the channel conditions that this user experiences, then more resources must be allocated to the user in order to achieve equal rate as other users. In other words, by using more resources, the user tries to compensate for lower SINR caused by the channel. This approach confirms the investigation in Chapter 3 in terms of the (un)suitability of a certain numerology for the environment. The second algorithm unites both, the optimal resources and numerology at the same time. This algorithm demonstrates that the previously used LTE transmission mode with a fixed 15 kHz subcarrier spacing performs well only in the situations when user velocity, and thereby the Doppler effect, is negligible; otherwise, it shows very poor performance despite the fact that it uses all available resources. A probing question is whether any other fixed numerology can achieve the same or approximately the same results as mixed numerology. The result depends on the environment being observed and the gain expected, that is, how willing one is to digress from the maximum gain (in case of mixed

numerology). Regarding the computational complexity of this algorithm, several ways to solve it are proposed. The algorithm is initially given by the optimal ILP solution. However, this kind of formulation can be too complex in dense environments or with a large bandwidth. Therefore, some less complex algorithms are proposed and the optimality provided by them is further discussed. The tradeoff of the accuracy of the results and the complexity is best achieved in the LP algorithm.

The algorithms described so far place emphasis on the eMBB use case proposed by 3GPP for 5G NR, where the main goal is to maximize the throughput. The uRLLC service is introduced as another use case in 5G NR for supporting stringent requirements in terms of latency and reliability. In order to meet the requirements of low latency imposed by the applications, previously developed algorithms explained in Chapter 5 are upgraded. The cost function is changed compared to the previous algorithms, with the goal to maximize the minimum achievable rate among NLLUs, ensuring, at the same time that the achievable rate of priority LLUs is large enough to transmit packets of a certain size within the required latency. This enables the comparison between two competing concepts, mixed numerology and mini-slots, under different channel and latency conditions. Which of the two approaches is preferable depends on the latency constraint as well as on how contradictory the channel conditions that users experience are. In the case of more flexible latency constraints, yet not conflicting channel conditions both methods perform similarly. In cases when the channel conditions are quite opposite, for example in indoor scenarios on the one hand, and high-speed environment on the other hand, the mixed numerology concept outperforms the mini-slot concept, since it pays off to employ more numerologies at the same time, even with a lack of resources assigned to the guard band. On the other hand, when the latency is stringent, the mixed numerology concept selects only the numerologies whose slot durations correspond to the latency constraint. Therefore, in such cases the mini-slot concept with the optimal subcarrier spacing outperforms the mixed numerology concept. The more general form would be to allow an interplay of mixed numerology and the mini-slot approach, that is seen as research to be considered in future. Additionally, UFMC in comparison to OFDM achieves only a minor gain thanks to less subcarriers assigned to the guard band. In order to reduce computational complexity for large-scale scenarios, the Dantzig-Wolfe decomposition method is applied, showing that it is possible to achieve the optimal solution with significantly reduced complexity.

7.2 Open Issues for Future Research

At the very end of the thesis, several possible suggestions are to be proposed for future research in order to improve the already proposed algorithms. Future research should consider some advanced scenarios, which are not the subject of this thesis, however, they certainly are worthy of closer inspection.

- Coexistence of mixed numerology and mini-slot concept - In the thesis, mixed numerology and mini-slots are observed separately, as two competing concepts. However, it is reasonable to look at these concepts together. It is expected that coexistence of these two approaches will allow for visibly better performance in the scenario of critical latency, since in such cases mixed numerology will have a complete set of numerologies at its disposal rather than just a narrow set. The superiority of mixed numerology over mini-slots can be seen under dissimilar channel conditions of users, when the use of multiple numerologies is necessary in order to maximize the performance. In all other cases, the gain may be negligible, given the fact that both approaches are already behaving similarly, leaving little possibility for enhancement.

- Optimal filter length and roll-off factor for minimizing INI - It was already emphasized that a longer filter improves the frequency localization and thus reduces INI on the one hand, but degrades the time localization, and thus increases the intra-numerology (intrinsic) interference on the other hand. The thesis examines the maximum filter length which disables the presence of the intrinsic interference. In addition to the optimal filter length, it is also desirable to examine the optimal roll-off factor and thereby to find a tradeoff between INI and intrinsic interference.

- Optimal numerology and resource allocation in multi-cell systems - The main challenge that arises with the introduction of a multi-cell system is the intercell interference that occurs when multiple cells attempt to use the same time-frequency resources or due to the misaligned resource blocks of different cells. This disables successful coordination among cells as well as formation of the mutually orthogonal reference signals across the cells. In order to prevent a misalignment among the cells, it is necessary to enable the placement of all numerologies in such a way that the interference is avoided.

- Applications of machine learning algorithms for large-scale scenarios - Since there is a plethora of requirements in terms of use cases, channel conditions as well as more heterogeneous structures of networks for beyond 5G generations, it is very important to investigate algorithms that would coexist with conventional algorithms to provide optimal solutions as efficiently as possible, or completely replace the conventional algorithms in situations where they cannot be applied [12, 81]. Machine learning algorithms are proposed as a possible solution. Their effectiveness is reflected in their ability to learn from data and easily establish a link between specific instances, without designing an application-based system. This becomes especially evident when allocating optimal resources and numerologies to users based on channel diversity conditions and QoS requirements.

Chapter 8

INI Derivation at Rx_1

Considering the OFDM system and assuming that the available bandwidth is equally shared by two users with different numerologies as shown in Fig. 8.1, in this section INI power caused by the user with a larger subcarrier spacing on the user with a smaller subcarrier spacing is derived.

After employing K_s-IDFT on the half of activated subcarriers of the *victim* user, the transmit signal can be expressed as:

$$\mathbf{x}^s = \left[x_0^{s,1}, x_1^{s,1}, \dots x_{K_s-1}^{s,1}\right]^T \in \mathbb{C}^{K_s \times 1}, s \in \mathcal{S}, \tag{8.1}$$

where

$$x_n^{s,1} = \frac{1}{\sqrt{K_s}} \sum_{k=0}^{\frac{1}{2}K_s} X_k^s e^{\frac{j2\pi nk}{K_s}}. \tag{8.2}$$

Given that $K_s = 2^{s'-1}K_{s'}$, within the same duration, the transmit signal of the aggressor user is composed of $2^{s'-1}$ symbols:

$$\begin{aligned} \mathbf{x}^{s'} &= \begin{bmatrix} \mathbf{x}^{s',1} \\ \vdots \\ \mathbf{x}^{s',2^{s'-1}} \end{bmatrix} \\ &= \begin{bmatrix} \left[x_0^{s',1}, x_1^{s',1}, \dots x_{K_{s'}-1}^{s',1}\right]^T \\ \vdots \\ \left[x_0^{s',2^{s'-1}}, x_1^{s',2^{s'-1}}, \dots x_{K_{s'}-1}^{s',2^{s'-1}}\right]^T \end{bmatrix} \in \mathbb{C}^{K_s \times 1}, s' \in \{2,3,4\}, s < s', \end{aligned} \tag{8.3}$$

where

$$x_n^{s',m} = \frac{1}{\sqrt{K_{s'}}} \sum_{k'=\frac{1}{2}K_{s'}}^{K_{s'}-1} X_{k'}^{s',m} e^{\frac{j2\pi nk'}{K_{s'}}}, \tag{8.4}$$

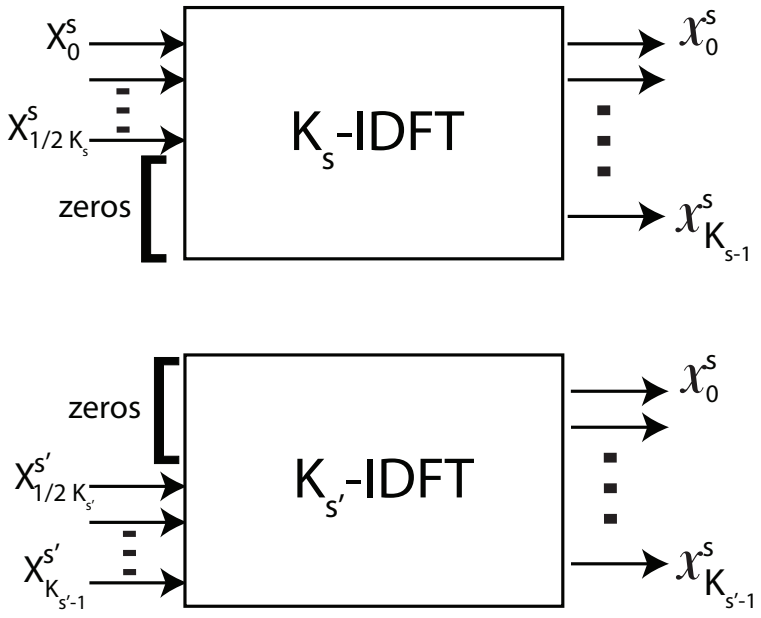


Figure 8.1: Simplified transmitter side of two numerologies.

with m being a symbol index and the total number of symbols is $2^{s'-1}$. Assuming a distortionless channel and noise-free transmission, the total received signal is the sum of the signals from the *victim* and *aggressor* users:

$$\mathbf{y} = \mathbf{x}^s + \mathbf{x}^{s'}, \;\; \mathbf{y} \in \mathbb{C}^{K_s \times 1}. \tag{8.5}$$

In the case when Rx_1 is applied, the estimated signal is given by:

$$\hat{\mathbf{X}}^s = [\hat{X}^s_0, \hat{X}^s_1, \cdots \hat{X}^s_{K_s-1}]^T = \mathfrak{F}\{\mathbf{y}\}, \tag{8.6}$$

where each of the estimated symbols is composed of the desired and interfering part caused by the *aggressor* user. At specific subcarrier l this is written as:

$$\hat{X}^s_l = \hat{X}^s_{\mathrm{des}_l} + \hat{X}^s_{\mathrm{INI}_l}, \tag{8.7}$$

The interference part can be further simplified into parts caused by specific symbols:

$$\begin{aligned}\hat{X}^s_{\mathrm{INI}_l} = \frac{1}{\sqrt{K_s}} \Bigg(& \sum_{n=0}^{K_{s'}-1} x^{s',1}_n e^{\frac{-j2\pi nl}{K_s}} + \cdots + \sum_{n=(m-1)K_{s'}}^{mK_{s'}-1} x^{s',m}_{n-(m-1)K_{s'}} e^{\frac{-j2\pi nl}{K_s}} \\ & + \sum_{n=(2^{s'-1}-1)K_{s'}}^{K_s-1} x^{s',2^{s'-1}}_{n-(2^{s'-1}-1)K_{s'}} e^{\frac{-j2\pi nl}{K_s}} \Bigg).\end{aligned} \tag{8.8}$$

In the following, initially only the interference caused by a single OFDM symbol m of the *aggressor* user is considered. For example, if the first sum is considered the interference is given by:

$$\hat{X}^{s}_{\mathrm{INI}_{l,1}} = \frac{1}{\sqrt{K_s}} \sum_{n=0}^{K'_s-1} x_n^{s',1} e^{\frac{-j2\pi nl}{K_s}}. \tag{8.9}$$

Substituting (8.4) in (8.9), interference at subcarrier l of the *victim* user can be written as:

$$\begin{aligned}
\hat{X}^{s}_{\mathrm{INI}_{l,1}} &= \frac{1}{\sqrt{K_{s'}K_s}} \sum_{n=0}^{K_{s'}-1} \sum_{k'=\frac{1}{2}K_{s'}}^{K_{s'}-1} X_{k'}^{s',1} e^{\frac{j2\pi nk'}{K_{s'}}} e^{\frac{-j2\pi nl}{K_s}} \\
&\text{substitute:} \quad k' = k + \frac{K_{s'}}{2} \\
&= \frac{1}{\sqrt{K_{s'}K_s}} \sum_{n'=0}^{K_{s'}-1} \sum_{k=0}^{\frac{1}{2}K_{s'}-1} X^{s',1}_{\left(\frac{1}{2}K_{s'}+k\right)} e^{\frac{j2\pi n(k+\frac{1}{2}K_{s'})}{K_{s'}}} e^{\frac{-j2\pi nl}{K_s}} \\
&\text{substitute:} \quad K_s = K_{s'}2^{s'-1} \\
&= \frac{1}{\sqrt{K_{s'}K_s}} \sum_{k=0}^{\frac{1}{2}K_{s'}-1} X^{s',1}_{\left(\frac{1}{2}K_{s'}+k\right)} \sum_{n=0}^{K_{s'}-1} e^{\frac{j2\pi n\left(k+\frac{1}{2}K_{s'}-\frac{l}{2^{s'-1}}\right)}{K_{s'}}}.
\end{aligned} \tag{8.10}$$

Furthermore, Euler's equation is utilized:

$$\begin{aligned}
P(\Omega) &= \sum_{n=0}^{K_{s'}-1} e^{j\Omega n} = \frac{1-e^{j\Omega K_{s'}}}{1-e^{j\Omega}} \\
&= \frac{\sin(\Omega K_{s'}/2)}{\sin(\Omega/2)} e^{j\Omega(K_{s'}-1)/2},
\end{aligned} \tag{8.11}$$

with $\Omega = \frac{2\pi}{K_{s'}}(k + \frac{1}{2}K_{s'} - \frac{l}{2^{s'-1}})$. By applying (8.11) on (8.10), $\hat{X}^{s}_{\mathrm{INI}_{l,1}}$ can be further simplified:

$$\hat{X}^{s}_{\mathrm{INI}_{l,1}} = \frac{1}{\sqrt{K_{s'}K_s}} \sum_{k=0}^{\frac{1}{2}K_{s'}-1} X^{s',1}_{\left(\frac{1}{2}K_{s'}+k\right)} \frac{\sin\left[\pi\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]}{\sin\left[\frac{\pi}{K_{s'}}\left(k - \frac{l}{2^{s'-1}} + \frac{1}{2}K_{s'}\right)\right]} e^{\frac{j\pi}{K_{s'}}\left(k-\frac{l}{2^{s'-1}}+\frac{1}{2}K_{s'}\right)(K_{s'}-1)}. \tag{8.12}$$

The interference power $\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l,1}}$ caused by a single OFDM symbol can be written as:

$$\begin{aligned}\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l,1}} &= \mathbb{E}[|\hat{X}^s_{\mathrm{INI}_{l,1}}|]^2 \\ &= \frac{1}{K_{s'}K_s}\sum_{k=0}^{\frac{1}{2}K_{s'}-1}\underbrace{\mathbb{E}\left[\left|X^{s',1}_{\left(\frac{1}{2}K_{s'}+k\right)}\right|\right]^2}_{=1}\frac{\sin^2\left[\pi\left(k-\frac{l}{2^{s'-1}}+\frac{1}{2}K_{s'}\right)\right]}{\sin^2\left[\frac{\pi}{K_{s'}}\left(k-\frac{l}{2^{s'-1}}+\frac{1}{2}K_{s'}\right)\right]} \\ &\cdot\underbrace{\mathbb{E}\left[\left|e^{\frac{j\pi}{K_{s'}}\left(k-\frac{l}{2^{s'-1}}+\frac{1}{2}K_{s'}\right)(K_{s'}-1)}\right|\right]^2}_{=1}.\end{aligned} \tag{8.13}$$

It is important to mention here that transmit symbols are mutually uncorrelated in time and frequency, i.e.:

$$\mathbb{E}[|X^m_k X^{m'}_{k'}|] = \delta(k-k')\delta(m-m'). \tag{8.14}$$

Taking into account total $2^{s'-1}$ OFDM symbols of the *aggressor* user with equal interference powers of all symbols, the total interference power $\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l}}$ experienced by the *victim* user is given by:

$$\begin{aligned}\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l}} &= 2^{s'-1}\sigma^{2,\mathrm{ofdm}}_{\mathrm{INI}_{\Delta f_s,l,1}} \\ &= \frac{1}{K^2_{s'}}\sum_{k=0}^{\frac{1}{2}K_{s'}-1}\frac{\sin^2\left[\pi\left(k-\frac{l}{2^{s'-1}}+\frac{1}{2}K_{s'}\right)\right]}{\sin^2\left[\frac{\pi}{K_{s'}}\left(k-\frac{l}{2^{s'-1}}+\frac{1}{2}K_{s'}\right)\right]}.\end{aligned} \tag{8.15}$$

Chapter 9

An Implementation of Dantzig-Wolfe Algorithm

In the following, all the necessary steps for the implementation of *Dantzig-Wolfe* algorithm are explained. The implementation is originally proposed by [82], but there are a few things that have to be adopted to the case considered in the thesis.

By providing a primal block angular structure represented by matrix $\mathbf{A}$ of dimensions $m \times n$ and depicted in Fig. 6.10, the input parameters to the algorithm are already defined: linking constraints $\mathbf{L}_1$ and $\mathbf{L}_2$, subproblem constraints related to NLLU and LLU, $\mathbf{A}_1$ and $\mathbf{A}_2$, respectively, as well as the number of the subblocks K, which in this case takes value $K = 2$:

$$\mathbf{A} = \begin{pmatrix} \mathbf{L}_1 & \mathbf{L}_2 \\ \mathbf{A}_1 & \\ & \mathbf{A}_2 \end{pmatrix}. \tag{9.1}$$

Submatrix A_k is of dimensions $m_k \times n_k$, where $k = 1, ...K$ and submatrix L_k is of dimensions $m_L \times n_k$. As already mentioned in Sec. 6.2.1, matrix $\mathbf{A}_1$ comprises all constraints relevant for NLLUs, that is, the first constraint of (6.3) relevant for NLLUs, as well as, (6.6b), (6.8b) and (6.9). Similarly, matrix $\mathbf{A}_2$ involves (6.4), (6.5) and (6.6a), i.e., the constraints relevant for LLUs. Linking submatrix L_k comprises constraints defined by (6.7) and (6.8a).

By considering such a structure, the linear program can be written as:

$$\begin{aligned} \text{min} \quad & \sum_{k=1}^{K} (\mathbf{c}^k)^T \mathbf{x}^k \\ \text{subject to} \quad & \sum_{k=1}^{K} \mathbf{L}_k \mathbf{x}^k \leq \mathbf{b}^0, \\ & \mathbf{A}^k \mathbf{x}^k \leq \mathbf{b}^k, \;\; k = \{1, 2\}, \\ & \mathbf{x}^k \geq 0, \;\; k = \{1, 2\}, \end{aligned} \tag{9.2}$$

where $\mathbf{x}^k$ represents the vector of variables that will be determined for a subproblem k and $\mathbf{c}^k$ and $\mathbf{b}^k$ are vectors of known coefficients, also specified for the subproblem k. Thus, the problem defined in 9.2 is composed of two parts:

- the master problem that is further composed of the linking (coupling) constraints represented by $\mathbf{L}_1$ and $\mathbf{L}_2$,

- two independent subproblems represented by $\mathbf{A}_1$ and $\mathbf{A}_2$.

The important feature of a subproblem k is that its feasible set can be represented as a polyhedron P_k. Any point on a polyhedron for the kth subproblem can be represented as a combination of extreme points denoted by $v_1^k, ..., v_{N_k}^k$ and extreme directions denoted by $d_1^k, ..., d_{1_k}^k$. Taking this into account and considering a master problem, the problem in 9.2 can be reformulated as:

$$\begin{aligned} \text{min} \quad & \sum_{k=1}^{K} (\mathbf{c}^k)^T \left(\sum_{i=1}^{N_k} \lambda_i^k v_i^k + \sum_{j=1}^{l_k} \mu_j^k d_j^k \right) \\ \text{subject to} \quad & \sum_{k=1}^{K} \mathbf{L}_k \left(\sum_{i=1}^{N_k} \lambda_i^k v_i^k + \sum_{j=1}^{l_k} \mu_j^k d_j^k \right) \leq \mathbf{b}^0 \\ & \sum_{i=1}^{N_k} \lambda_i^k = 1, \;\; k = \{1, 2\}, i = \{1, ..., N_k\}, \\ & \sum_{j=1}^{l_k} \mu_j^k = 1, \;\; k \in \{1, 2\}, , j = \{1, ..., l_k\}, \\ & \lambda_i^k \geq 0, \mu_j^k \geq 0. \end{aligned} \tag{9.3}$$

Since the number of λ_i^k and μ_j^k can be extremely large, the idea is to restrict the master problem by applying a revised simplex method. Since the observed problem is bounded, μ_j^k does not exist and

9.3 can be further simplified:

$$\begin{aligned} \text{min} \quad & \sum_{k=1}^{K} (\mathbf{c}^k)^T \sum_{i=1}^{N_k} \lambda_i^k v_i^k \\ \text{subject to} \quad & \sum_{k=1}^{K} \mathbf{L}_k \sum_{i=1}^{N_k} \lambda_i^k v_i^k + \mathbf{s} = \mathbf{b}^0, \\ & \sum_{i=1}^{N_k} \lambda_i^k = 1, \;\; k = \{1,2\}, i = \{1,...,N_k\}, \\ & \lambda_i^k \geq 0. \end{aligned} \tag{9.4}$$

Having this in mind, in the following step-by-step implementation is explained.
Step 0 implies an initialization of the basis $\mathbf{B}$ and the basic variables $\mathbf{x_B}$ for the restricted master problem. It is important to mention that the initial basic variables associated to the master problem have to take the values from the feasible region, and that does not have to be zero. The basis has the following shape:

$$\mathbf{B} = \begin{pmatrix} \mathbf{I}_{m_k \times m_k} & (\mathbf{L}_1 \mathbf{v}^1)_{m_L \times 1} & (\mathbf{L}_2 \mathbf{v}^2)_{m_L \times 1} \\ \mathbf{0}_{K \times m_k} & \mathbf{I}_{K \times K} & \end{pmatrix}, \tag{9.5}$$

and

$$\mathbf{x_B} = \begin{pmatrix} \mathbf{b}^0_{m_L \times 1} - (\mathbf{L}_1 \mathbf{v}^1)_{m_L \times 1} - (\mathbf{L}_2 \mathbf{v}^2)_{m_L \times 1} \\ \mathbf{I}_{K \times 1} \end{pmatrix} = \begin{pmatrix} \mathbf{s}_{m_L \times 1} \\ \mathbf{I}_{K \times 1} \end{pmatrix}. \tag{9.6}$$

Step 1 refers to the calculation of the dual variables π of the restricted master problem based on which a suitable objective function of the subproblems is defined, i.e., $\pi = \mathbf{f}_B^T \mathbf{B}^{-1}$:

$$\begin{pmatrix} \pi^1 \\ \pi_1^2 \\ \pi_2^2 \end{pmatrix} = \mathbf{f}_B^T \mathbf{B}^{-1}, \tag{9.7}$$

where $\mathbf{f}_B$ is a vector composed of the product $(\mathbf{c}^k)^T v_i^k$ related to a variable λ_i^k, π^1 are dual variables associated with the linking constraints, π_1^2 are dual variables associated with the first subproblem, which is further related to NLLUs and π_2^2 are dual variables associated with the second subproblem, related to LLUs.

In *Step* 2, the optimization of the subproblems is executed simultaneously in parallel, since they are independent. The optimization for the subproblem k can be formulated as:

$$\begin{aligned} \min \quad & \sigma_k = \Big((\mathbf{c}^k)^T - (\pi^1)^T \mathbf{L}_k\Big)\mathbf{x}^k \\ \text{subject to} \quad & \mathbf{A}_k \mathbf{x}^k \leq \mathbf{b}^k, \\ & \mathbf{x}^k \geq 0. \end{aligned} \tag{9.8}$$

According to the obtained optimal extreme point $\mathbf{x}^k = \mathbf{v}_i^k*$ the reduced cost is calculated:

$$r_*^k = \sigma_k^* - \pi_k^2. \tag{9.9}$$

A condition that checks whether the reduced cost of the basic feasible solution is non-negative value or not is then defined. In a case it is non-negative value, the revised simplex method stops with an optimal solution:

$$r_{min} = \min_{k \in \{1,..,K\}} \{r_*^k\} \geq 0. \tag{9.10}$$

As long as the reduced cost is the negative value and the problem is bounded, a column generation operation is performed, i.e., the column that is associated to the optimal extreme point v_i^t* is generated, where t is an index of k in subproblem k, such that $r_{min} = r_*^t$ and that column enters the basis:

$$\bar{\mathbf{a}} = \begin{pmatrix} \mathbf{L}_k v_i^t* \\ \mathbf{e}_t \end{pmatrix}. \tag{9.11}$$

In a case that all subproblems are unbounded, an extreme direction has to be generated. In the thesis it is not the case.

In *Step* 4, a descent direction $\mathbf{d}$ is defined and thereby it is declared whether a linear problem is unbounded or not:

$$\mathbf{B}\mathbf{d} = -\bar{\mathbf{a}}. \tag{9.12}$$

In this case it is bounded, i.e., $\mathbf{d} < 0$, the next step is executed.

In *Step* 5 the step length based on the minimum ratio test is calculated:

$$\alpha = \min_{l \in \bar{\mathbf{B}}} \Big\{ -\frac{\mathbf{x}_l}{\mathbf{d}_l} | \mathbf{d}_l < 0 \Big\}, \tag{9.13}$$

with $\bar{\mathbf{B}}$ denoting the vector of the basic variables.

Based on the step length, a basic vector is recalculated in *Step* 6:

$$\mathbf{x}_B = \mathbf{x}_B + \alpha \mathbf{d}. \tag{9.14}$$

Finally, in *Step* 7 a basis $\mathbf{B}$ is updated by adding a column $\bar{\mathbf{a}}$ and the column B_l* is discarded, where l_* is the index of basic variable that achieves the minimum ratio test. The process continues until a reduced cost takes a non-negative value, i.e., there is no column that enters the basis.

The output parameters of the optimization process are: the optimal solution, objective function at that value and the binary variable that describes an exit condition, i.e., whether an optimal solution is found or LP is unbounded.

Chapter 10

List of Abbreviations

ACLR Adjacent Channel Leakage power Ratio

AMPS Advanced Mobile Phone System

AWGN Additive White Gaussian Noise

BER bit-error ratio

CP Cyclic Prefix

DFT Discrete Fourier Transform

DFTS Discrete Fourier Transform-Spread

eMBB Enhanced Mobile Broadband

FBMC Filtered Bank Multicarrier

FDM Frequency Division Multiplexing

FIR Finite Impulse Response

f-OFDM filtered OFDM

GB Guard Band

GFDM Generalized Frequency Division Multiplexing

GSM Global System for Mobile Communications

IBI Interband Interference

ICI Intercarrier Interference

IDFT Inverse Discrete Fourier Transform

ILP Integer Linear Programming

INI Internumerology Interference

ISI Intersymbol Interference

LL Low Latency

LP Linear Programming

LS Least Square

LTE Long Term Evolution

MIMO Multiple Input Multiple Output

mMTC Massive Machine-Type Communications

MSE Mean Squared Error

NR New Radio

OBUE Operating Band Unwanted Emissions

OFDM Orthogonal Frequency Division Multiplexing

OOB Out-of-Band

PAPR Peak-to-Average-Power-Ratio

PDP Power Delay Profile

PHY Physical Layer

QoS Quality of Service

RB Resource Block

RE Resource Element

RMS Root Mean Square

SINR Signal to Interference plus Noise Ratio

SMS Short Message Service

SNR Signal to Noise Ratio

TDL Tapped Delay Line

UFMC Universal Filtered Multicarrier

UMTS Universal Mobile Telecommunications System

uRLLC Ultra-Reliable Low Latency Communications

WOLA Weighted Overlap-and-Add

WSSUS Wide Sense Stationary Uncorrelated Scattering

ZP Zero Padding

1G first generation

2G second generation

3G third generation

3GPP Third Generation Partnership Project

4G fourth generation

5G fifth generation

Bibliography

[1] A. Zaidi, F. Athley, J. Medbo, U. Gustavsson, G. Durisi, and X. Chen, *5G Physical Layer: principles, models and technology components*. Academic Press, 2018.

[2] A. A. Zaidi, R. Baldemair, H. Tullberg, H. Bjorkegren, L. Sundstrom, J. Medbo, C. Kilinc, and I. Da Silva, "Waveform and numerology to support 5G services and requirements", *IEEE Communications Magazine*, vol. 54, no. 11, pp. 90–98, 2016.

[3] 3. G. P. P. (3GPP), "Study on scenarios and requirements for next generation access technologies", 3rd Generation Partnership Project (3GPP), TR 38.913, May 2017.

[4] M. Series, "IMT Vision–Framework and overall objectives of the future development of IMT for 2020 and beyond", *ITU Recommendation*, pp. 2083–, 2015.

[5] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions", *IEEE Communications Magazine*, vol. 54, no. 9, pp. 59–65, 2016.

[6] M. Simsek, A. Aijaz, M. Dohler, J. Sachs, and G. Fettweis, "5G-enabled tactile internet", *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 3, pp. 460–473, 2016.

[7] M. Shafi, A. F. Molisch, P. J. Smith, T. Haustein, P. Zhu, P. De Silva, F. Tufvesson, A. Benjebbour, and G. Wunder, "5G: A tutorial overview of standards, trials, challenges, deployment, and practice", *IEEE journal on selected areas in communications*, vol. 35, no. 6, pp. 1201–1221, 2017.

[8] 3. G. P. P. (3GPP), "Study on New Radio (NR) access techology", 3rd Generation Partnership Project (3GPP), TR 38.912, Jun. 2017.

[9] P. Guan, D. Wu, T. Tian, J. Zhou, X. Zhang, L. Gu, A. Benjebbour, M. Iwabuchi, and Y. Kishiyama, "5G field trials: OFDM-based waveforms and mixed numerologies", *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1234–1243, 2017.

[10] Z. E. Ankarali, B. Peköz, and H. Arslan, "Flexible Radio Access beyond 5G: A future projection on waveform, numerology, and frame design principles", *IEEE Access*, vol. 5, pp. 18 295–18 309, 2017.

[11] 3. G. P. P. (3GPP), "Technical specification group services and system aspects; Release 15 Description; Summary of Rel-15 work items", TR 21.915, 2019.

[12] A. Yazar and H. Arslan, "A waveform parameter assignment framework for 6G with the role of machine learning", *IEEE Open Journal of Vehicular Technology*, vol. 1, pp. 156–172, 2020.

[13] L. You, Q. Liao, N. Pappas, and D. Yuan, “Resource optimization with flexible numerology and frame structure for heterogeneous services”, *IEEE Communications Letters*, vol. 22, no. 12, pp. 2579–2582, 2018.

[14] A. González, S. Kühlmorgen, A. Festag, and G. Fettweis, “Resource allocation for block-based multi-carrier systems considering QoS requirements”, in *IEEE Global Communications Conference - GLOBECOM*, IEEE, 2017, pp. 1–7.

[15] L. Zhang, A. Ijaz, P. Xiao, A. Quddus, and R. Tafazolli, “Subband filtered multi-carrier systems for multi-service wireless communications”, *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1893–1907, 2017.

[16] H. Ji, S. Park, J. Yeo, Y. Kim, J. Lee, and B. Shim, “Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects”, *IEEE Wireless Communications*, vol. 25, no. 3, pp. 124–130, 2018.

[17] G. Durisi, T. Koch, and P. Popovski, “Toward massive, ultrareliable, and low-latency wireless communication with short packets”, *Proceedings of the IEEE*, vol. 104, no. 9, pp. 1711–1726, 2016.

[18] S. Schwarz, T. Philosof, and M. Rupp, “Signal processing challenges in cellular-assisted vehicular communications: Efforts and developments within 3GPP LTE and beyond”, *IEEE Signal Processing Magazine*, vol. 34, no. 2, pp. 47–59, 2017.

[19] S. Schwarz and M. Rupp, “Society in motion: Challenges for LTE and beyond mobile communications”, *IEEE Communications Magazine*, vol. 54, no. 5, pp. 76–83, 2016.

[20] N. Patriciello, S. Lagen, L. Giupponi, and B. Bojovic, “5G New Radio numerologies and their impact on the end-to-end latency”, in *IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, IEEE, 2018, pp. 1–6.

[21] C. Campolo, A. Molinaro, F. Romeo, A. Bazzi, and A. O. Berthet, “5G NR V2X: On the impact of a flexible numerology on the autonomous sidelink mode”, in *IEEE 2nd 5G World Forum (5GWF)*, IEEE, 2019, pp. 102–107.

[22] Z. Zhang, Y. Gao, Y. Liu, and Z. Li, “Performance evaluation of shortened transmission time interval in LTE networks”, in *IEEE Wireless Communications and Networking Conference (WCNC)*, IEEE, 2018, pp. 1–5.

[23] L. Marijanović, S. Schwarz, and M. Rupp, “Optimal numerology in OFDM systems based on imperfect channel knowledge”, in *87th Vehicular Technology Conference (VTC Spring)*, IEEE, 2018, pp. 1–5.

[24] ——, “Pilot pattern optimization for small data packet transmission”, in *26th Telecommunications Forum (TELFOR)*, IEEE, 2018, pp. 1–4.

[25] S. Pratschner, B. Tahir, L. Marijanović, M. Mussbah, K. Kirev, R. Nissel, S. Schwarz, and M. Rupp, “Versatile mobile communications simulation: The Vienna 5G Link Level Simulator”, *EURASIP Journal on Wireless Communications and Networking*, 2018.

[26] S. Schwarz, B. R. Elbal, E. Zöchmann, L. Marijanović, and S. Pratschner, "Dependable wireless connectivity: Insights and methods for 5G and beyond", *e & i Elektrotechnik und Informationstechnik*, vol. 135, no. 7, pp. 449–455, 2018.

[27] L. Marijanović, S. Schwarz, and M. Rupp, "Intercarrier interference of multiple access UFMC with flexible subcarrier spacings", in *25th European Signal Processing Conference (EUSIPCO), 2017*, IEEE, 2017, pp. 888–892.

[28] ——, "Optimal resource allocation with flexible numerology", in *IEEE International Conference on Communication Systems (ICCS)*, IEEE, 2018, pp. 136–141.

[29] ——, "A novel optimization method for resource allocation based on mixed numerology", in *IEEE International Conference on Communications (ICC) 2019*, IEEE, 2019, pp. 1–6.

[30] ——, "Multi-user resource allocation for low latency communications based on mixed numerology", in *90th Vehicular Technology Conference (VTC2019-Fall)*, IEEE, 2019, pp. 1–7.

[31] ——, "Multiplexing services in 5G and beyond: Optimal resource allocation based on mixed numerology and mini-slots", vol. 8, 2020, pp. 209 537–209 555.

[32] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. Ten Brink, I. Gaspar, N. Michailow, A. Festag, *et al.*, "5GNOW: Non-orthogonal, asynchronous waveforms for future mobile applications", *IEEE Communications Magazine*, vol. 52, no. 2, pp. 97–105, 2014.

[33] B. Yang, X. Zhang, L. Zhang, A. Farhang, P. Xiao, and M. Ali Imran, "Windowed OFDM for mixed-numerology 5G and beyond systems", *Radio Access Network Slicing and Virtualization for 5G Vertical Industries*, pp. 43–61, 2020.

[34] I. Gaspar, L. Mendes, M. Matthé, N. Michailow, A. Festag, and G. Fettweis, "LTE-compatible 5G PHY based on generalized frequency division multiplexing", in *11th International Symposium on Wireless Communications Systems (ISWCS)*, IEEE, 2014, pp. 209–213.

[35] M. Schellmann, Z. Zhao, H. Lin, P. Siohan, N. Rajatheva, V. Luecken, and A. Ishaque, "FBMC-based air interface for 5G mobile: Challenges and proposed solutions", in *9th international conference on cognitive radio oriented wireless networks and communications (CROWNCOM)*, IEEE, 2014, pp. 102–107.

[36] R. Zayani, Y. Medjahdi, H. Shaiek, and D. Roviras, "WOLA-OFDM: A potential candidate for asynchronous 5G", in *IEEE Globecom Workshops (GC Wkshps)*, IEEE, 2016, pp. 1–5.

[37] Y. Liu, X. Chen, Z. Zhong, B. Ai, D. Miao, Z. Zhao, J. Sun, Y. Teng, and H. Guan, "Waveform candidates for 5G networks: Analysis and comparison", *arXiv preprint arXiv:1609.02427*, 2016.

[38] H. Lin and P. Siohan, "Major 5G waveform candidates: Overview and comparison", *Signal Processing for 5G: Algorithms and Implementations*, 2016.

[39] R. Gerzaguet, N. Bartzoudis, L. G. Baltar, V. Berg, J.-B. Doré, D. Kténas, O. Font-Bach, X. Mestre, M. Payaró, M. Färber, *et al.*, "The 5G candidate waveform race: A comparison of complexity and performance", *EURASIP Journal on Wireless Communications and Networking*, vol. 2017, no. 1, p. 13, 2017.

[40] F. Schaich and T. Wild, "Waveform contenders for 5G? OFDM vs. FBMC vs. UFMC", in *6th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, IEEE, 2014, pp. 457–460.

[41] T. Wild, F. Schaich, and Y. Chen, "5G air interface design based on universal filtered (UF-) OFDM", in *19th International Conference on Digital Signal Processing*, IEEE, 2014, pp. 699–704.

[42] N. Michailow, M. Matthé, I. S. Gaspar, A. N. Caldevilla, L. L. Mendes, A. Festag, and G. Fettweis, "Generalized frequency division multiplexing for 5th generation cellular networks", *IEEE Transactions on Communications*, vol. 62, no. 9, pp. 3045–3061, 2014.

[43] R. Nissel, "Filter bank multicarrier modulation for future wireless systems", Ph.D. dissertation, Dissertation, TU Wien, 2017.

[44] T. Strohmer and S. Beaver, "Optimal OFDM design for time-frequency dispersive channels", *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1111–1122, 2003.

[45] G. Kongara, C. He, L. Yang, and J. Armstrong, "A comparison of CP-OFDM, PCC-OFDM and UFMC for 5G uplink communications", *IEEE Access*, vol. 7, pp. 157 574–157 594, 2019.

[46] Y. Medjahdi, Y. Louet, M. B. Mabrouk, D. Roviras, R. Zakaria, H. Shaiek, D. L. Ruyet, S. Traverso, R. Gerzaguet, D. Demmer, J. Dore, and R. Zayani, "Impact of selective channels on post-OFDM waveforms for 5G machine type communications", in *15th International Symposium on Wireless Communication Systems (ISWCS)*, 2018, pp. 1–5.

[47] P. Banelli, S. Buzzi, G. Colavolpe, A. Modenini, F. Rusek, and A. Ugolini, "Modulation formats and waveforms for 5G networks: Who Will Be the Heir of OFDM?: An overview of alternative modulation schemes for improved spectral efficiency", *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 80–93, 2014.

[48] K. Mahender, T. A. Kumar, and K. Ramesh, "PAPR analysis of fifth generation multiple access waveforms for advanced wireless communication", *International Journal of Engineering & Technology*, vol. 7, no. 3.34, pp. 487–490, 2018.

[49] P. N. Rani and C. S. Rani, "UFMC: The 5G modulation technique", in *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, 2016, pp. 1–3.

[50] M. N. Tipan, J. Caceres, M. N. Jimenez, I. N. Cano, and G. Arevalo, "Comparison of clipping techniques for PAPR reduction in UFMC systems", in *IEEE 9th Latin-American Conference on Communications (LATINCOM)*, 2017, pp. 1–4.

[51] "R1-165014, Subband-wise filtered OFDM for new radio below 6 GHz", *3GPP TSG-RAN WG1 85*, May 2016.

[52] S. Wang, J. S. Thompson, and P. M. Grant, "Closed-form expressions for ICI/ISI in filtered OFDM systems for asynchronous 5G uplink", *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4886–4898, 2017.

[53] J. Wen, J. Hua, W. Lu, Y. Zhang, and D. Wang, "Design of waveform shaping filter in the UFMC system", *IEEE Access*, vol. 6, pp. 32 300–32 309, 2018.

[54] S. Geng, X. Xiong, L. Cheng, X. Zhao, and B. Huang, "UFMC system performance analysis for discrete narrow-band private networks", in *IEEE 6th International Symposium on Microwave, Antenna, Propagation, and EMC Technologies (MAPE)*, IEEE, 2015, pp. 303–307.

[55] L. Zhang, A. Ijaz, P. Xiao, M. A. Imran, and R. Tafazolli, "MU-UFMC system performance analysis and optimal filter length and zero padding length design", *arXiv preprint arXiv:1603.09169*, 2016.

[56] L. Zhang, P. Xiao, and A. Quddus, "Cyclic prefix-based universal filtered multicarrier system and performance analysis", *IEEE Signal Processing Letters*, vol. 23, no. 9, pp. 1197–1201, 2016.

[57] 3. G. P. P. (3GPP), "NR; Physical channels and modulation", 3rd Generation Partnership Project (3GPP), TS 38.211, 2018.

[58] Y. Li and L. J. Cimini, "Bounds on the interchannel interference of OFDM in time-varying impairments", *IEEE Transactions on Communications*, vol. 49, no. 3, pp. 401–404, 2001.

[59] H.-P. Kuchenbecker *et al.*, "Intercarrier and intersymbol interference analysis of OFDM systems on time-varying channels", in *4th IEEE Workshop on Signal Processing Advances in Wireless Communications. SPAWC 2003*, IEEE, 2003, pp. 140–144.

[60] M. Šimko, P. S. Diniz, Q. Wang, and M. Rupp, "Adaptive pilot-symbol patterns for MIMO OFDM systems", *IEEE Transactions on Wireless Communications*, vol. 12, no. 9, pp. 4705–4715, 2013.

[61] R. Nissel, M. Lerch, M. Simko, and M. Rupp, "Bit error probability for pilot-symbol-aided OFDM channel estimation in doubly-selective channels", in *WSA 2014; 18th International ITG Workshop on Smart Antennas*, VDE, 2014, pp. 1–6.

[62] G. Matz and F. Hlawatsch, "Fundamentals of time-varying communication channels", in *Wireless Communications Over Rapidly Time-Varying Channels*, Elsevier, 2011, pp. 1–63.

[63] X. Wang, T. Wild, and F. Schaich, "Filter optimization for carrier-frequency-and timing-offset in universal filtered multi-carrier systems", in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, IEEE, 2015, pp. 1–6.

[64] X. Wang and S. t. Brink, "SINR analysis of different OFDM-inspired waveforms over doubly dispersive channels", *IEEE Transactions on Vehicular Technology*, vol. 69, no. 9, pp. 9459–9468, 2020.

[65] X. Wang, T. Wild, F. Schaich, and S. ten Brink, "Pilot-aided channel estimation for universal filtered multi-carrier", in *IEEE 82nd Vehicular Technology Conference (VTC2015-Fall)*, IEEE, 2015, pp. 1–5.

[66] 3. G. P. P. (3GPP), "Study on channel model for frequencies from 0.5 to 100 GHz", 3rd Generation Partnership Project (3GPP), TR 38.901, Jan. 2018.

[67] R. M. Rao, V. Marojevic, and J. H. Reed, "Adaptive pilot patterns for CA-OFDM systems in nonstationary wireless channels", *IEEE Transactions on Vehicular Technology*, vol. 67, no. 2, pp. 1231–1244, 2017.

[68] J.-W. Choi and Y.-H. Lee, "Optimum pilot pattern for channel estimation in OFDM systems", *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2083–2088, 2005.

[69] ITU, "Recommendation ITU-R M.1225: Guidelines for Evaluation of Radio Transmission Technologies for IMT-2000", *International Telecommunication Union, Tech. Rep*, 1997.

[70] A. B. Kihero, M. S. J. Solaija, A. Yazar, and H. Arslan, "Inter-numerology interference analysis for 5G and beyond", in *IEEE Globecom Workshops (GC Wkshps)*, 2018, pp. 1–6.

[71] P. Guan, D. Wu, T. Tian, J. Zhou, X. Zhang, L. Gu, A. Benjebbour, M. Iwabuchi, and Y. Kishiyama, "5G field trials: OFDM-Based waveforms and mixed numerologies", *IEEE Journal on Selected Areas in Communications*, vol. 35, no. 6, pp. 1234–1243, 2017.

[72] J. Mao, L. Zhang, P. Xiao, and K. Nikitopoulos, "Interference analysis and power allocation in the presence of mixed numerologies", *IEEE Transactions on Wireless Communications*, vol. 19, no. 8, pp. 5188–5203, 2020.

[73] L. Zhang, A. Ijaz, P. Xiao, K. Wang, D. Qiao, and M. A. Imran, "Optimal filter length and zero padding length design for universal filtered multi-carrier (UFMC) system", *IEEE Access*, vol. 7, pp. 21 687–21 701, 2019.

[74] 3. G. P. P. (3GPP), "5G; NR; Base Station (BS) radio transmission and reception", 3rd Generation Partnership Project (3GPP), TS 38.104, 2018.

[75] ——, "5G; Study on New Radio (NR) access techology", TR 38.912, 2018.

[76] T. Pinto, C. Alves, R. Mansi, and J. Valério de Carvalho, "Solving the multiscenario max-min Knapsack problem exactly with column generation and branch-and-bound", *Mathematical Problems in Engineering*, vol. 2015, 2015.

[77] 3. G. P. P. (3GPP), "5G; NR;User Equipment (UE) radio transmission and reception; Part 1: Range 1 Standalone", TS 38.101, 2018.

[78] J. R. Tebboth, "A computational study of Dantzig-Wolfe decomposition", *PhD thesis, University of Buckingham*, 2001.

[79] F. Vanderbeck, "On Dantzig-Wolfe decomposition in integer programming and ways to perform branching in a branch-and-price algorithm", *Operations Research*, vol. 48, no. 1, pp. 111–128, 2000.

[80] J. van den Brand, "A deterministic linear program solver in current matrix multiplication time", *arXiv:1910.11957v2 [cs.DS]*, 2020.

[81] J. Zhang, X. Xu, K. Zhang, B. Zhang, X. Tao, and P. Zhang, "Machine learning based flexible transmission time interval scheduling for eMBB and uRLLC coexistence scenario", *IEEE Access*, vol. 7, pp. 65 811–65 820, 2019.

[82] R. H. Kwon, *Introduction to Linear Optimization and Extensions with MATLAB®*. CRC Press, 2013.