Faculty of Computer Science, Electrical Engineering and Mathematics
Computer Networks Group

Enhancing Cellular Handovers: Optimizing Quality of Experience using Machine Learning

Master's Thesis

in Partial Fulfillment of the Requirements for the
Degree of
Master of Science

by
Björn Luchterhandt

advised by
Marvin Illian

submitted to
Prof. Dr. Lin Wang
and
Dr. Simon Oberthür

Paderborn, February 26, 2025

Disclaimer

I hereby declare that I prepared this thesis entirely on my own and have not used outside sources without declaration in the text. Any concepts or quotations applicable to these sources are clearly attributed to them. This thesis has not been submitted in the same or substantially similar version, not even in part, to any other authority for grading and has not been published elsewhere.

Paderborn, February 26, 2025

Björn Luchterhandt

Abstract

Cellular handovers ensure the continuous connectivity of mobile devices as they move through various mobile network cells. Handover decisions in mobile networks can significantly impact the Quality of Experience of mobile device users as they influence both the achievable data rate and the power consumption of a mobile device. This thesis optimizes cellular handovers using a machine learning technique (Reinforcement Learning), considering both power consumption and data rates. To quantify power consumption of mobile devices, a power model is developed and validated with power measurements in a commercial mobile network. The model is subsequently integrated into a system-level simulator for cellular handovers and applied to jointly optimize data rates and power consumption of mobile network devices with respect to Quality of Experience. Both the developed model and the handover simulation environment are publicly available and can be applied and extended in future network research.

Acknowledgements

The power measurements in this work have been supported by Deutsche Telekom AG. I would like to thank the team at Deutsche Telekom for their support and cooperation – gaining insights into the configuration of radio access networks has been a fascinating experience. I would further like to thank everyone who has supported my research through ideas, perspectives, and resources: Marvin Illian for always having a sympathetic ear and offering a second perspective when I was unsure about the best approach. The Computer Networks Group for the friendly and supportive environment, a workplace, coffee, and a 5G modem for the power measurements. Tim Hetkämper and the Measurement Engineering Group for providing me with hardware, know-how, and an awesome Python library to abstract the low-level hardware communication out of the way. The E-Lab of the Electrical Engineering Student Council for borrowing me their hardware. Prof. Dr. Thomas Richthammer from the Stochastics working group in the Department of Mathematics, who spontaneously took the time to offer his extensive expertise. Finally, the team at actiVita Paderborn for offering their gym as a measurement location, even for over-night measurements – your christmas punch has made the measurements far more enjoyable.

Thank you, all!

Contents

List of Figures

List of Tables

1 Introduction

Over the past decade, traffic demands on mobile radio networks have been rapidly increasing [12]. Not only is the number of mobile subscribers steadily growing, but with 5G deployments on the rise, the average per-device mobile data rate is also increasing significantly. Network operators are striving to achieve high customer satisfaction by providing high data rates, oftentimes deploying more mobile network infrastructure. However, recent studies have shown that there are unutilized data rate gains in many existing deployments due to a suboptimal assignment of mobile devices to radio cells [13], [14].

A cell is uniquely identifiable via a broadcasted cell ID and is powered by a single base station on a specific carrier frequency [15], meaning one base station can provide cells on various frequencies, and cells from various base stations can share the same carrier frequency. As a mobile device, subsequently referred to as User Equipment (UE), is moving, it has to switch between various cells to maintain connectivity. The choice of cells heavily impacts a UE's achievable data rate due to varying cell bandwidths, traffic loads, and signal strengths.

Since UEs are typically battery-powered and the technological enhancements in battery energy densities do not keep up with the increasing demands [16], UE power consumption is an important aspect of mobile radio networks too. The power consumed for cellular data transmissions depends on a variety of factors, including transmission power and radio bandwidth. From a UE's perspective, these parameters can wildly vary for different cells due to diverse radio channel conditions. Hence, cell selection plays a major role in power consumption of mobile devices.

The 3rd Generation Partnership Project (3GPP) differentiates between cell reselection and handovers [17], [18]. Cell reselection is performed by UEs in the Radio Resource Control (RRC) idle state, i.e., when no RRC connection is established, and denotes the UE-initiated process of selecting another cell for future connections [17]. Reselection follows a 3GPP-specified algorithm which measures Reference Signal Received Power (RSRP) for available cells and autonomously decides which cell to use, based on parameters broadcasted by the base station. A handover, on the other hand, is initiated by the Radio Access Network (RAN) and hands an existing RRC connection over to another cell while a UE is in RRC connected mode [18]. In order to make informed handover decisions, the RAN can configure UEs to collect measurements about alternative cells in RRC connected mode and report them back to the base station [5].

Cell reselection and handovers complement each other to ensure connectivity throughout all RRC states. However, they differ fundamentally in the amount of control a network

operator has about their outcomes. Because the cell reselection algorithm is provided by the 3GPP specification and implemented in the UE's modem chip, a RAN can only tune the parameters it provides to it, without customizing the algorithm itself. Handovers, on the other hand, are RAN-initiated and network operators can freely implement custom handover decision logic in their networks. Hence, they provide a mechanism to optimize cell selection for various metrics without requiring changes to the 3GPP specification.

Mobile operators strive to maximize users' perceived Quality of Experience (QoE), i.e., how satisfied a user is with the provided network service. Another common optimization metric in mobile networking research is the data rate that a UE can achieve [14], [19], [20], [21]. Yet, it has been shown that data rates do not linearly impact QoE [22], [23]. Therefore, maximizing UE data rates in a cellular network does not necessarily optimize the QoE of individual users. Moreover, a user's QoE is also impacted by UE battery life [24] which, in turn, depends on power consumption. Since handover decisions affect both data rates and UE power consumption, optimizing handover decisions can improve QoE. A thorough literature review yielded no existing approaches for QoE-driven multi-metric handover optimization with respect to data rates and UE power consumption. This thesis therefore focuses on optimizing handovers to improve QoE, based on data rates and UE power consumption.

A popular approach to handover optimization is the use of Reinforcement Learning (RL) [20], [25], which is able to dynamically adapt handover decisions to the observed system behavior. Considering the complex dynamic environment at hand, this is preferable over static expert-designed algorithms which can be challenging to develop and hard to extend for additional metrics in future work. The work in this thesis hence consists of two major parts:

(1) The design and validation of a power model to estimate UE power consumption
(2) The training and evaluation of an RL agent for handover decisions, including the development of a system-level simulator for 5G handovers

In the following sections, an introduction is given to relevant concepts and mechanisms in 5G (Section 1.1) and the core concepts of RL (Section 1.2). Chapter 2 presents related work in both fields of this thesis: Handover optimization and power modeling. Based on this, Chapter 3 focuses on the design and validation of a power model, and Chapter 4 deals with handover optimization using RL. Finally, the results of this thesis are discussed in Chapter 5, and a conclusion is drawn in Chapter 6.

1.1 An Introduction to 5G New Radio

The 5th Generation (5G) mobile network, standardized by the 3GPP, enables significantly higher data rates than its predecessor Long-Term Evolution (LTE) and supports a wider frequency spectrum [26]. These changes to the physical layer have been introduced via a new radio access technology, 5G New Radio (NR) [1]. 5G NR's physical layer is built around many of the concepts of LTE's physical layer, but enhancing it with more flexible configuration options [26], [27]. Therefore, if NR and LTE cells are synchronized, they can be operated alongside each other on the same frequency range without interfering with each other.

5G also specifies a new core network, providing a gradual adoption path [1] as illustrated in Figure 1. Perspectively, 5G Base Stations (BSs) are designed to be operated with a 5G core network. This mode of operation is called 5G Standalone (5G SA). However, 5G BSs (and therefore 5G NR) can also be used in conjunction with an LTE core network to offer more bandwidth in an existing LTE deployment by combining LTE and NR carrier frequencies. This mode of operation is called 5G Non-Standalone (5G NSA).

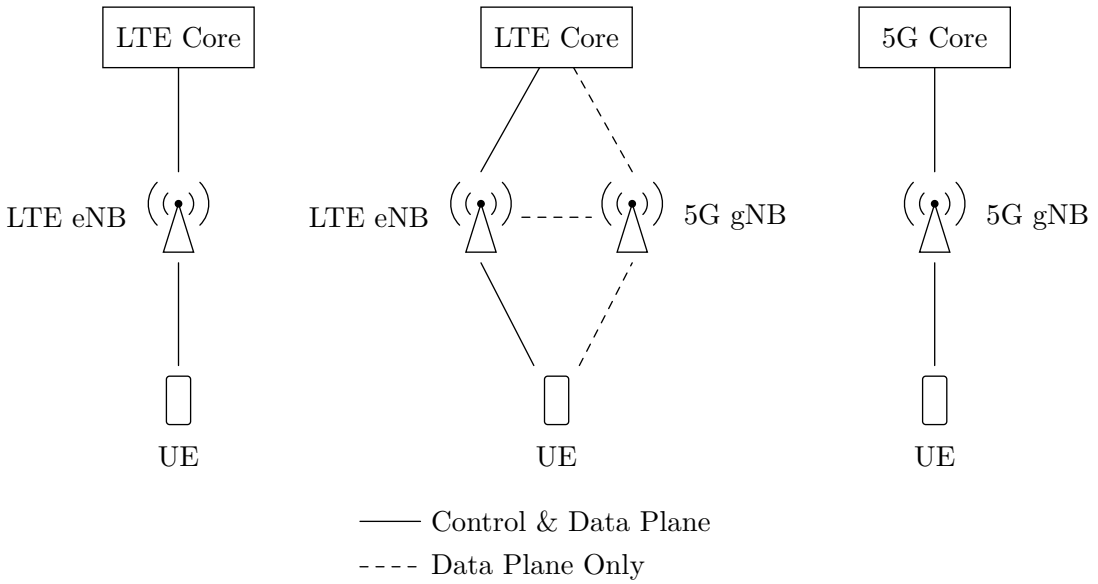


Figure 1: LTE vs 5G NSA vs 5G SA Architectures (left-to-right), based on [1]

Developing a 5G UE power model requires a detailed understanding of the processes involved in 5G NR communication to model the UE's activity and power states. The following sections therefore cover the basic functionality of 5G NR by elaborating on the time and frequency domains (Sections 1.1.1 and 1.1.2, respectively), physical channel concepts and operation (Section 1.1.3), and common power saving mechanisms (Section 1.1.4).

1.1.1 Time Domain and Numerologies

5G NR utilizes Orthogonal Frequency-Division Multiplexing (OFDM), both for Downlink (DL) and Uplink (UL) transmissions [9]. This translates to transmitting multiple symbols in parallel on multiple subcarrier frequencies. These subcarrier frequencies are distributed evenly across the desired frequency spectrum. While the DL in LTE relies on OFDM too, its subcarrier frequencies are always (with a single exception) spaced 15 kHz apart from each other [1]. By contrast, NR makes the Subcarrier Spacing (SCS) configurable based on an integer μ, referred to as *Numerology* [9]:

$$\text{SCS} = 2^{\mu} \cdot 15 \text{ kHz} \tag{1}$$

Based on the numerology, the NR time domain is structured via a number of units (depicted in Figure 2), as follows. The largest unit is a *Frame*, denoting a duration of 10 ms. One frame contains 10 *Subframes*, where each subframe is one millisecond long. A subframe, in turn, consists of a number of *Slots*. The precise number of slots in a subframe depends on the numerology. Namely, for a given numerology μ, each subframe comprises 2^{μ} slots. For instance, at 15 kHz SCS ($\mu = 0$), one subframe contains one slot, while at 60 kHz SCS ($\mu = 2$), a subframe contains four slots. Finally, a slot allows for 14 OFDM symbols to be transmitted (optionally only 12 OFDM symbols for $\mu = 2$) [9]. Transmitting one OFDM symbol, in this context, means transmitting one symbol on each of a set of OFDM subcarrier frequencies.

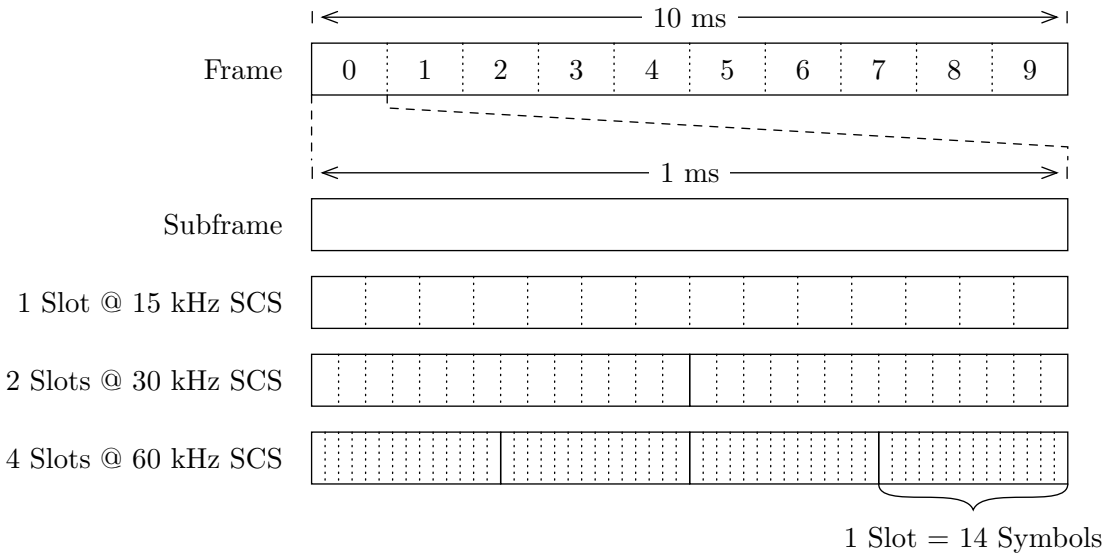


Figure 2: 5G New Radio Frame Structure

Complementing Figure 2, the correlation between numerology, SCS, and slots is listed in Table 1. In practice, the prevalent numerology in mid-band 5G deployments is $\mu = 1$, i.e., 30 kHz SCS [28].

Table 1: Impact of the Numerology on SCS and Slots [9]

μ	SCS	Slots per Subframe	Slots per Frame	Slot Duration
0	15 kHz	1	10	1 ms
1	30 kHz	2	20	0.5 ms
2	60 kHz	4	40	0.25 ms
3	120 kHz	8	80	0.125 ms

5G NR supports both Time-Division Duplexing (TDD) – where UL and DL transmissions share the same frequency range, but do not occur at the same time – and Frequency-Division Duplexing (FDD), where UL and DL transmissions occur simultaneously using distinct frequency ranges. Which mode is used depends on the authority-defined properties of the allocated spectrum (paired or unpaired spectrum) [1].

When TDD is applied, the RAN configures slot formats, dictating which symbols of a slot are reserved for DL transmissions, UL transmissions, or can be used for either of both [2]. As demonstrated in Figure 3, a slot can be configured as entirely DL, entirely UL, or a combination of both, defined by a number of DL symbols at the beginning of the slot, a number of subsequent flexible symbols, and a number of UL symbols at the end of the slot [2].

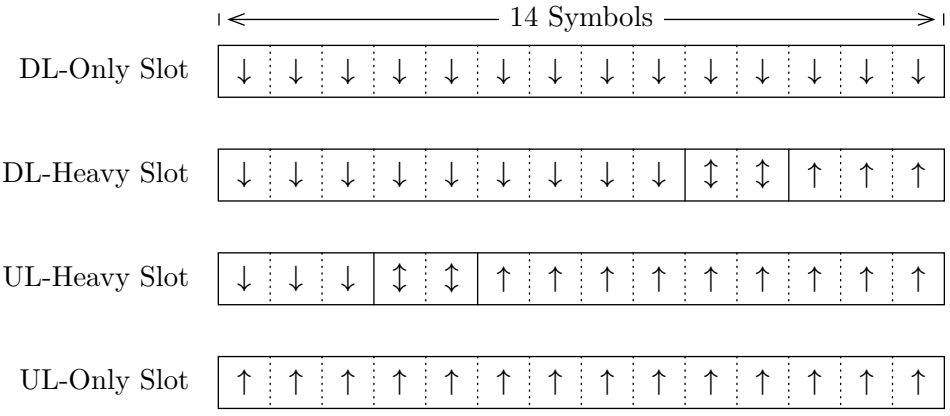


Figure 3: Example Slot Formats [1], [2]

1.1.2 Frequency Domain and Resource Grid

NR can be operated on a wider range of frequencies than LTE, most notably adding a new Frequency Range (FR), *FR2*, for millimeter-wave communication [26], as illustrated

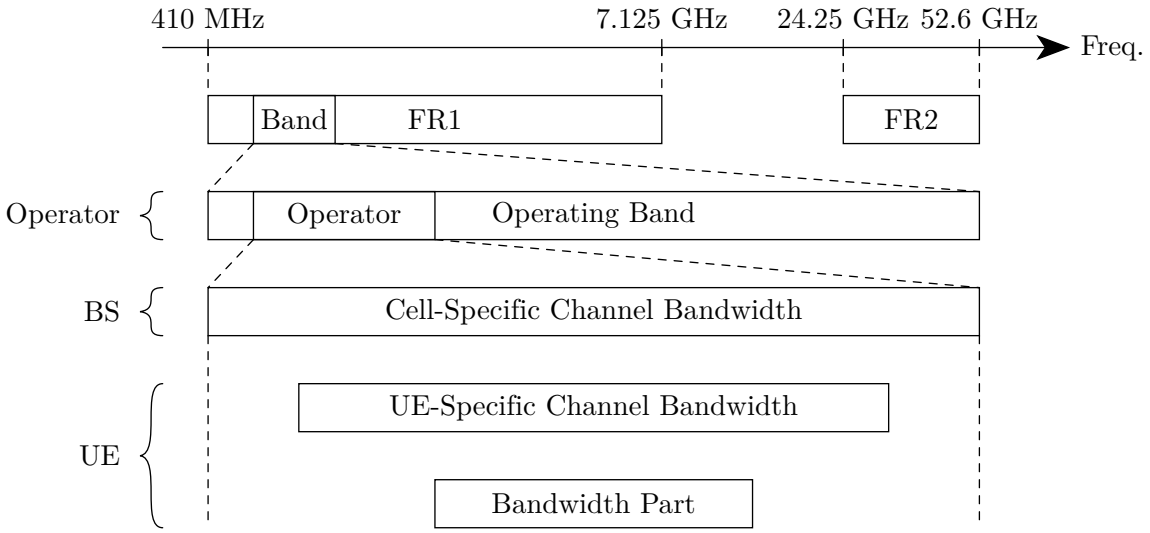


Figure 4: Spectrum Usage of 5G New Radio [3]

on the frequency axis in Figure 4. Within a FR, 3GPP specifies multiple *Operating Bands*, portions of which may be assigned to specific mobile operators [3]. A mobile operator, in turn, may use its portion of an operating band to deploy cells on various frequencies and with various bandwidths. In the EU, the predominant 5G band is n78, i.e., TDD with frequencies in the range of 3.3 to 3.8 GHz and a cell bandwidth of at most 100 MHz [28].

Some UEs only support limited bandwidths, which is why a dedicated UE-specific channel bandwidth is included in Figure 4 [3]. Moreover, within the UE-specific channel bandwidth, the RAN can configure the UE to only use specific Bandwidth Parts (BWPs) [3]. UL and DL BWPs are configured separately, but need to be centered around the same frequency in TDD mode [3].

Combining the physical resources in the time and frequency domains, the smallest allocatable resource unit is one Resource Element (RE), i.e., the bandwidth of one subcarrier for the duration of one symbol [9]. When multiple REs are visualized in the time and frequency domains, the individual REs compose a grid structure, the *Resource Grid*, as depicted in Figure 5. Moreover, in the frequency domain, a bandwidth of 12 subcarriers is denoted as a Resource Block (RB) [9].

1.1.3 Physical Channels and Physical/MAC Layer Operation

5G NR builds upon the concept of physical channels [9]. A physical channel can be thought of as a set of REs used for a designated function. BSs and UEs adhere to 3GPP-defined rules to determine which physical channels they transmit on or decode at a given time, i.e., how resource elements are mapped to sending and receiving devices. This section provides an overview of NR's physical channels, their purposes, and their interplay in the physical

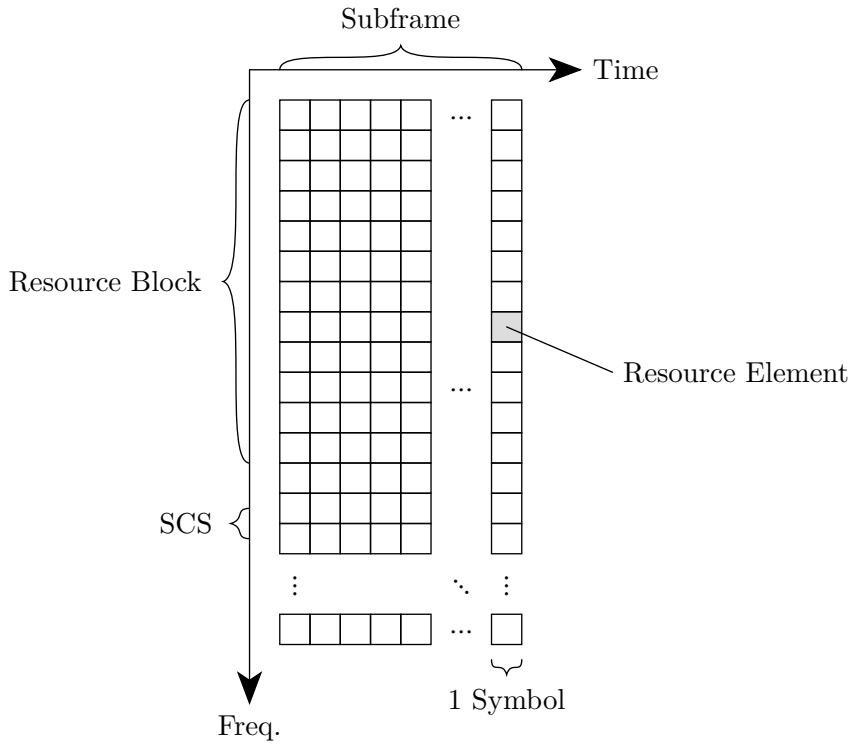


Figure 5: 5G NR Resource Grid Concept

and Medium Access Control (MAC) layers to serve UEs with wireless network access. All physical channels are listed in Table 2, intended to serve as a reference in the remainder of this thesis. In the following paragraphs, the channels and their interaction will be detailed based on a practical example, namely, the process of initial synchronization and RRC connection establishment of a UE with a BS, outlined in Figure 6.

In order for a UE to be able to connect to a cell, the UE needs to gather basic information about the cell, such as the timing and numerology. For this purpose, each cell periodically broadcasts a Signal Synchronization Block (SSB) using a group ("block") of neighboring REs [2]. The SSB contains a reference signal that the UE also uses to measure a cell's Reference Signal Received Power (RSRP). Once a UE is synchronized to a cell, it can decode basic information from it, which is broadcasted on the Physical Broadcast Channel (PBCH). The PBCH comprises a periodic block of REs known as the Master Information Block (MIB). A MIB is sent alongside each SSB and indicates the cell's numerology, current system frame number, and other details required for further interaction with the cell [5].

With the MIB decoded, the UE can access the cell's Physical Downlink Shared Channel (PDSCH), which carries DL payload transmissions and a number of System Information Blocks (SIBs) for all UEs. To further interact with the cell, the UE then needs to decode the periodically transmitted SIB1 from the PDSCH [4]. Among other information, SIB1

Table 2: 5G NR Physical Channels [9]

Channel	Purpose
Physical Broadcast Channel (PBCH)	Broadcasting the MIB to UEs
Physical Downlink Control Channel (PDCCH)	Announcing resource assignments (DCI) to UEs
Physical Downlink Shared Channel (PDSCH)	• Transmitting payload data to UEs • Broadcasting System Information Blocks (SIBs) to UEs
Physical Random Access Channel (PRACH)	Initiating cell access
Physical Uplink Control Channel (PUCCH)	• Sending Scheduling Requests (SRs) for UL data • Acknowledging DL data • Sending Channel State Information (CSI) reports
Physical Uplink Shared Channel (PUSCH)	Transmitting payload data to BS

lists the resources allocated for the Physical Random Access Channel (PRACH), allowing the UE to initially request cell access by transmitting a *Random Access Preamble* on the PRACH.

The next steps require a basic understanding of how the BS announces DL and UL transmissions. Central to this is the Physical Downlink Control Channel (PDCCH) [9]. It is transmitted by the BS and received by all active UEs in a cell. The PDCCH carries Downlink Control Information (DCI) messages, which announce the scheduling decisions of the BS, i.e., which UE may transmit which REs to the BS and which UE shall receive and decode which REs from the BS. A single DCI message only announces scheduling decisions to one UE at a time. Thus, multiple DCI messages may need to be transmitted in parallel on the PDCCH, requiring a mechanism to multiplex DCI messages. NR solves this by applying Code-Division Multiple Access (CDMA) [29]. Each UE has a cell-unique ID, the Radio Network Temporary Identifier (RNTI), which is used to generate a pseudo-random bit sequence for CDMA, known to both the UE and the BS. The DCI message signal is then "scrambled" (XORed) with the bit sequence by the BS, transmitted in parallel with all other DCIs on the PDCCH, and finally retained at the UE again by applying the bit sequence to the received signal [29]. In addition to the resource assignments for a UE, a DCI message also dictates the Modulation and Coding Scheme (MCS) for DL and UL transmissions.[1]

[1] Available modulation techniques are Quadrature Phase Shift Keying (QPSK) and 16/64/256 Quadrature Amplitude Modulation (QAM) [29].

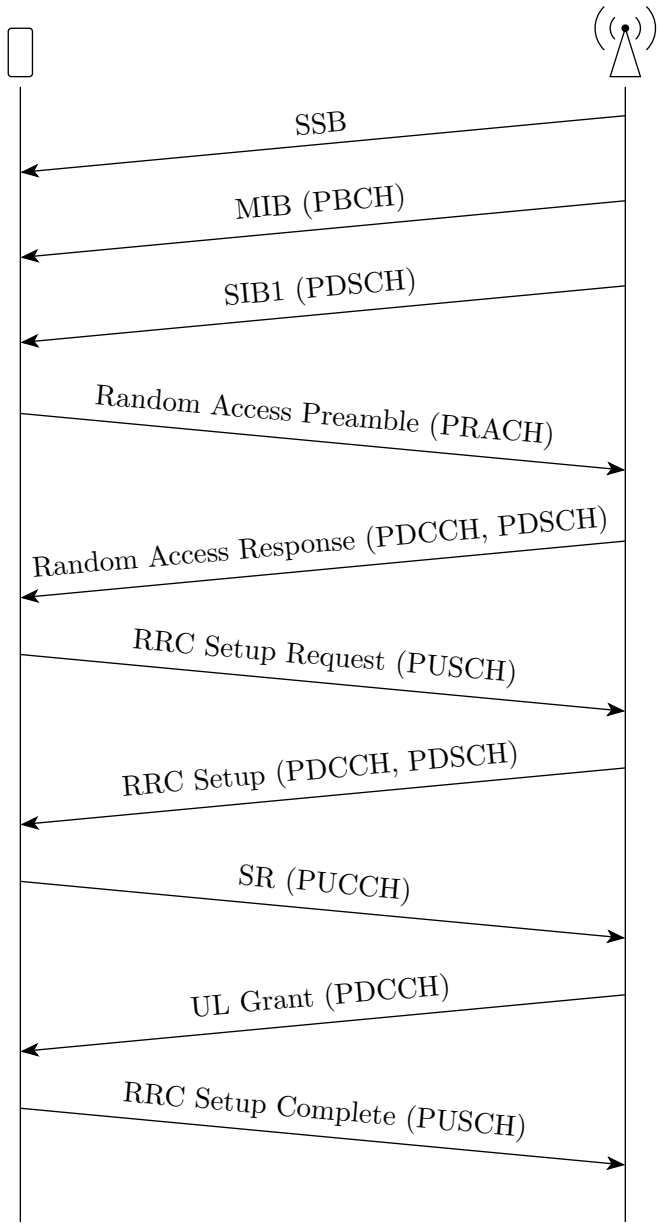


Figure 6: Synchronization and RRC Connection Sequence [4], [5]

Upon receiving the UE's Random Access Preamble, the BS replies with a *Random Access Response*.[2] Like other DL transmissions, the Random Access Response is first announced via a DCI on the PDCCH and subsequently transmitted on the PDSCH [2]. The random access procedure has completed when the UE receives the Random Access Response. At this point, the BS is aware of the UE and takes responsibility for scheduling resources for its UL and DL transmissions. Therefore, further UL messages are transmitted on the Physical Uplink Shared Channel (PUSCH), and the PRACH is no longer used by the UE.

[2]This description assumes that the random access preamble does not collide with another UE's access attempt at the BS. The collision handling can be considered irrelevant within the scope of this thesis.

Before the UE can send and receive payload data, it needs to establish an RRC connection by sending an *RRC Setup Request*. The BS has already included a PUSCH resource assignment ("uplink grant") for the RRC Setup Request transmission in its Random Access Response [2]. Once the BS has received the RRC Setup Request, the RAN handles the connection establishment and – upon success – announces the RRC connection to the UE in an *RRC Setup* message [5]. Like all further DL transmissions, the transmission of the RRC Setup message is announced on the PDCCH and carried out on the PDSCH [2].

The final step for establishing an RRC connection is an acknowledgement from the UE via an *RRC Setup Complete* message to the BS [5]. This involves a channel which has not previously been used up to this point, the Physical Uplink Control Channel (PUCCH) [9]. The PUCCH provides a mechanism for UEs to transmit a fixed set of small control messages to the BS without explicit scheduling by the BS. Most notably, this is required for Scheduling Requests (SRs). A SR is sent by a UE on the PUCCH to indicate that it has buffered UL data to be transmitted on the PUSCH and is awaiting an UL grant from the BS [2]. Messages on the PUCCH are multiplexed using frequency hopping, i.e., while transmitting a PUCCH message, the UE switches between different subcarriers according to a unique pattern known by both the UE and the BS [9]. This pattern is provided to the UE in the PUCCH configuration. Because it is unique for each UE in a cell, the PUCCH configuration is communicated to each UE individually in the RRC Setup message [5]. In the previous paragraph, it was mentioned that the BS includes an UL grant for the RRC Setup Request in its Random Access Response. This is necessary because the UE sends the RRC Setup Request prior to receiving the RRC Setup message with the PUCCH configuration from the BS. Therefore, the PUCCH configuration is not yet available to the UE when sending the RRC Setup Request, and the UE cannot send a SR for the transmission. By contrast, when acknowledging the RRC setup with an RRC Setup Complete message, the UE already has PUCCH access and thus uses a SR to receive an UL grant for transmitting the RRC Setup Complete message [2]. The BS, in turn, reacts to the SR by scheduling PUSCH resources for the RRC Setup Complete message and communicating them to the UE in a DCI via the PDCCH. Once the UE receives the UL grant, it transmits the RRC Setup Complete message on the PUSCH, completing the RRC setup upon reception at the BS.

In addition to the basic functionality described in the previous paragraphs, the physical and MAC layers also provide a number of mechanisms that have not been mentioned in

the given example, but are nonetheless relevant in the context of this thesis. The following list briefly presents these mechanisms.

- **MIMO and Beamforming**: 5G NR supports Multiple Input, Multiple Output (MIMO), a technology which exploits spacial diversity by using multiple UE Receive (Rx) and Transmit (Tx) antennas in parallel to increase bandwidth utilization [30]. This is combined with *beamforming* at the BS. Beamforming uses an array of antennas to concentrate the intensity of individual radio signals ("beams") in specific directions, based on constructive and destructive interference between phase-shifted, amplitude-scaled radio signals transmitted on several antennas in parallel [30]. Notably, the BS transmits SSBs as a burst of beams in various directions, allowing the UE to report its preferred beam to the BS [2], [4]. The BS can use this information to adapt its beam for PDSCH transmissions to the UE, referred to as a "link beam" in some literature [20].
- **CSI Measurements and Reporting**: The BS periodically transmits a Channel State Information Reference Signal (CSI-RS) which the UE uses to gather information about the radio conditions and report them back to the BS [31]. The reporting conditions, frequency, and which metrics the UE reports are determined by the RAN and provided to the UE via RRC messages [31]. Measurement reports can be transmitted via the PUCCH or the PUSCH [31].
- **Sounding Reference Signals (SRSs)**: SRSs are periodically transmitted by connected UEs based on RRC-configured settings provided by the RAN [9]. They complement CSI reports by allowing the BS to perform detailed channel estimation measurements itself, as opposed to receiving measurement results from the UE.
- **RRM Measurements and Reporting**: Similar to CSI measurements, the RAN can configure Radio Resource Management (RRM) measurements [5]. While CSI measurements consider the serving cell, RRM measurements gather information about neighboring cells to enable the RAN to make informed handover decisions. RRM measurements can be triggered on specified conditions, e.g., if the serving cell RSRP drops below a configured threshold [5].
- **HARQ**: Hybrid Automatic Repeat Requests is the mechanism responsible for acknowledging transmissions and requesting retransmissions in case of transmission errors. HARQ utilizes forward error correction in combination with positive (ACK) and negative (NACK) acknowledgements based on checksums [29]. Acknowledgements are typically piggy-backed on PUSCH/PDSCH transmissions [2]. Otherwise, they can also be transmitted individually on the PUCCH or included in the DCI on the PDCCH, depending on the transmission direction [2].

- **PUCCH Formats**: There are PUCCH formats with various lengths and purposes. Short PUCCH formats only require up to two symbols and can be used for SRs or HARQ acknowledgements [26]. Longer PUCCH formarts are available to increase coverage by transmitting more symbols for the same amount of payload information [26].
- **Semi-Persistent Scheduling**: Instead of announcing each PDSCH transmission individually via a DCI message on the PDCCH, the RAN can also configure a recurrent schedule via RRC messages [32]. DCI is then used to announce the activation or deactivation of a given schedule. A similar mechanism is also available for the PUSCH, where it is called *Configured Grant* [32].

1.1.4 Power Saving Mechanisms

Because UEs are typically battery-powered, 5G NR has been designed with a focus on energy efficiency. This involves several mechanisms to reduce the power consumption of UEs, most notably Discontinuous Reception (DRX) and BWP Switching. Both techniques will be presented in the following two sections.

1.1.4.1 Discontinuous Reception

DRX exploits the fact that a large amount of UE energy is consumed by frequently monitoring the PDCCH, checking wheter the BS transmits a DCI message to inform the UE of an upcoming DL transmission. However, mobile network user traffic is oftentimes irregular and inherently bursty [33]. This may result in long inter-arrival times between traffic bursts, during which the UE monitors the PDCCH without any DL traffic being announced by the BS. DRX allows the UE to enter a sleep state when no traffic is available, only periodically waking up to be informed if DL traffic has arrived at the BS in the meantime [32]. Depending on the user traffic pattern, this can significantly reduce UE power consumption at the cost of higher maximum latencies. To avoid disruptive latencies, the DRX parameters for a UE thus have to be selected in alignment with the UE's Quality of Service (QoS) requirements. 5G tracks QoS requirements via 5G QoS Identifiers (5QIs) [34]. The standard 5QI for web traffic (buffered video streaming, TCP-based web apps, etc.) dictates a packet delay budget of 300 ms [34], i.e., the DRX parameters need to be set s.t. a UE never sleeps more than 300 ms at a time.

5G NR defines two types of DRX: Connected Mode DRX (C-DRX) and Idle Mode DRX. C-DRX applies when an RRC connection is active and, as described above, allows the UE to temporarily enter a sleep state instead of monitoring the PDCCH [32]. Idle Mode DRX, on the other hand, applies when no RRC connection is established and controls the periodic reception of paging messages from a BS [34]. As the focus of this thesis is on

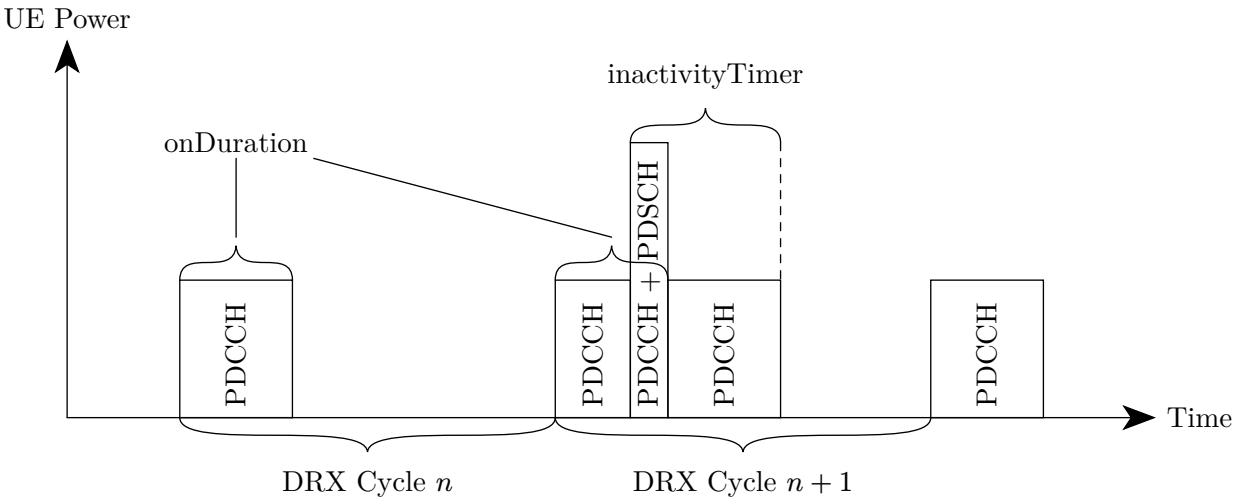


Figure 7: UE Activity over Several DRX Cycles

handover optimization, which does not involve the RRC Idle state, only C-DRX will be considered in the remainder of this work. As such, the terms "DRX" and "C-DRX" will be used synonymously. The following paragraphs further explain the C-DRX mechanism and its parameters.

The DRX concept is based on cycles with a configurable cycle duration and starting time (subframe) [32]. Figure 7 depicts the activity of a UE over several DRX cycles. The UE keeps track of various timers, most importantly an *On duration timer* and an *inactivity timer*.[3] Timers are set to RRC-configured values on specific events and trigger actions when they expire. At the beginning of each cycle (or a configured number of slots after the cycle's beginning), the UE sets its On duration timer and monitors the PDCCH while the timer is running. This time period is also referred to as the *On duration*. If no transmission has been announced during the On duration, the UE enters a sleep state when the On duration timer expires. If any DL packet arrives at the BS while the UE is asleep, the BS buffers the packet and transmits it at the beginning of the next cycle's On duration.

Additionally, whenever a transmission (DL or UL) is announced on the PDCCH, the UE (re)sets its inactivity timer. While the inactivity timer is running, the UE keeps monitoring the PDCCH, even if the On duration timer has already expired. If the inactivity timer expires and the On duration timer is not active, the UE enters a sleep state again. In Figure 7, no packet arrives during the On duration of DRX cycle n, allowing the UE to sleep for the rest of the cycle. At the beginning of the On duration in DRX cycle n+1, there is no PDSCH transmission. It can therefore be assumed that no DL packet has arrived at

[3]In addition to these timers, the DRX specification also defines timers for nested ("short") DRX cycles, as well as HARQ retransmission timers for UL and DL [32]. Those details will not be considered here, given that their impact on UE power consumption is presumably small in relation to the added complexity.

the BS while the UE was asleep. Shortly before the end of the On duration in cycle n+1, the BS schedules a PDSCH transmission, making the UE set its inactivity timer and stay active despite the expiry of the On duration timer. In this example, there is no follow-up transmission and the inactivity timer expires after a while, allowing the UE to enter a sleep state for the rest of cycle n+1. Conceptually, the inactivity timer ensures that a UE stays active for the entire duration of a traffic burst and only becomes inactive between bursts. Therefore, DRX does not disrupt active data flows, but saves UE energy when no data is being transmitted.

The example in Figure 7 only involves DL traffic. Yet, traffic bursts may also be initiated by UL transmissions, for instance when a UE requests a web page. In that case, there is no need for the UE to wait for the next On duration. Instead, when the UE receives UL data from upper network layers while in a DRX sleep period, it immediately wakes up and transmits a SR to the BS, subsequently monitoring the PDCCH for the corresponding UL grant [32]. Upon receiving the UL grant, the UE sets its inactivity timer and proceeds as usual. Figure 8 demonstrates the UE activities when UL traffic arrives in a DRX sleep period. It is worth noting that the inactivity timer is set when receiving the DCI for a transmission, not upon the transmission itself. This causes a gap between the start of the inactivity timer and the PUSCH transmission in Figure 8, given that PUSCH transmissions are announced several slots in advance [31].

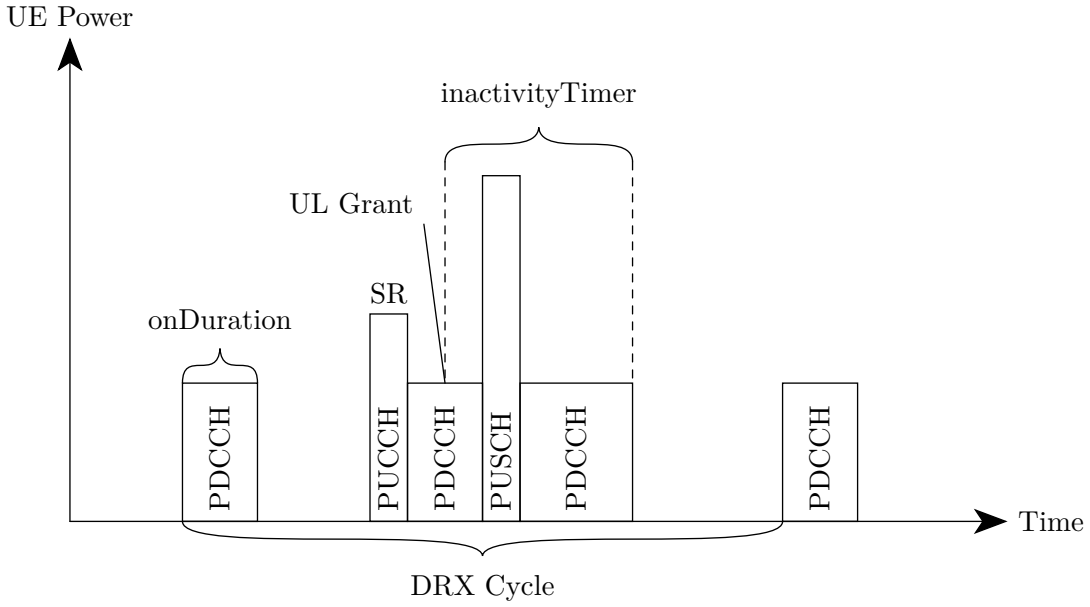


Figure 8: UL Traffic Handling in a DRX Sleep Period

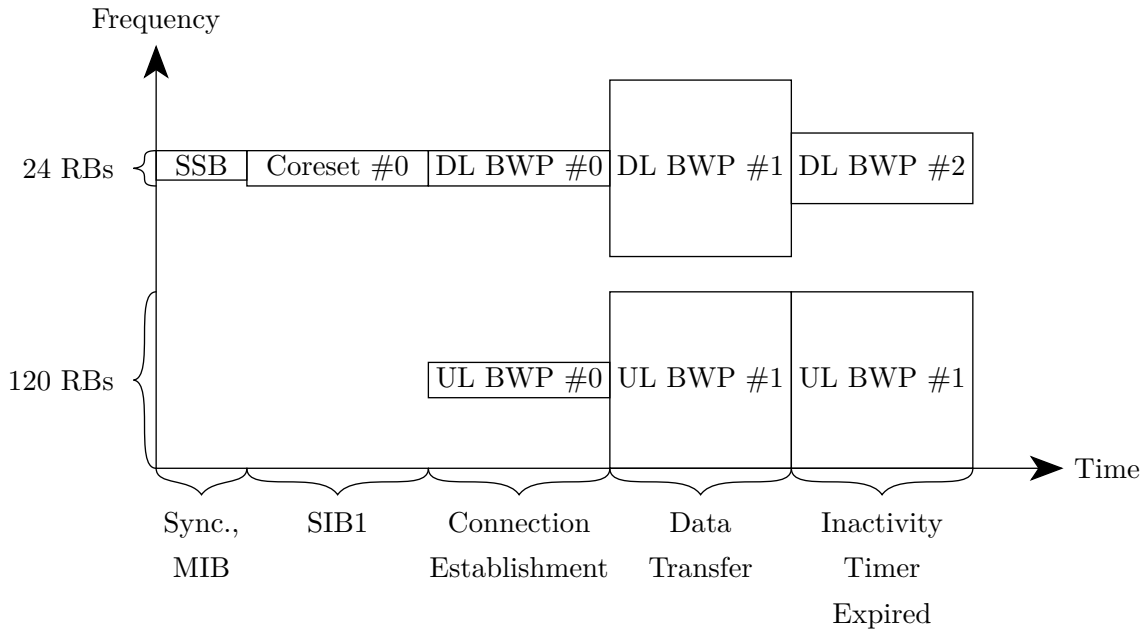


Figure 9: BWP Switching Example [3]

1.1.4.2 BWP Switching

The power that a UE requires to monitor the PDCCH is rouqhly proportional to the bandwidth across which the PDCCH is received [10]. Given the large cell bandwidths that 5G NR supports, considerable amounts of UE energy can thus be saved by reducing the bandwidth for a UE when it does not utilize its full bandwidth. BWP switching denotes the practice of dynamically selecting the DL or UL BWPs used by a UE [3]. It can either be performed on behalf of the RAN via RRC signalling, or automatically based on an inactivity timer.

Figure 9 demonstrates BWP switching for a UE which connects to a NR cell (operated in FDD mode) and transfers a burst of payload data, followed by an idle period. Initially, during synchronization and connection establishment, only a small BWP (the "initial" BWP) is used. The initial BWP is communicated to the UE in SIB1 [3]. Once the RRC connection has been set up, however, the DL and UL BWPs are switched via RRC signalling, providing a larger bandwidth for the payload data transfer. In addition to RRC-signalled BWP switching, the UE in this example has also been configured with an inactivity timer, similar to the DRX inactivity timer. Once the inactivity timer expires, the UE switches to the configured default BWP [3]. In Figure 9, this mechanism is applied to the DL BWP to reduce the power required for PDCCH monitoring when there is no transmission activity. The size of the UL BWP only impacts the power that the UE consumes for UL transmissions (including occasional SRSs) and therefore has no significant

effect on UE power consumption when there is no payload data. Therefore, the UL BWP is not switched by the inactivity timer in this example.

1.2 Reinforcement Learning

The previous section has provided an introduction to 5G NR, detailing the relevant concepts and mechanisms built upon in this thesis. Similarly, this section introduces Reinforcement Learning (RL). While RL has evolved into a broad field of research over the past decades [35], the foundations required for its application in this thesis can be summarized much more concisely than is the case with 5G NR.

RL is a field of Machine Learning in which an agent learns to solve a task by interacting with an environment [35]. More specifically, as illustrated in Figure 10, an RL agent repeatedly selects an action from a provided action space and communicates it to the environment. For each action taken, the agent receives an observation of the environment's state, as well as a numeric reward signal. The goal of the agent is to choose its actions in such a way that the reward signal is maximized [35].

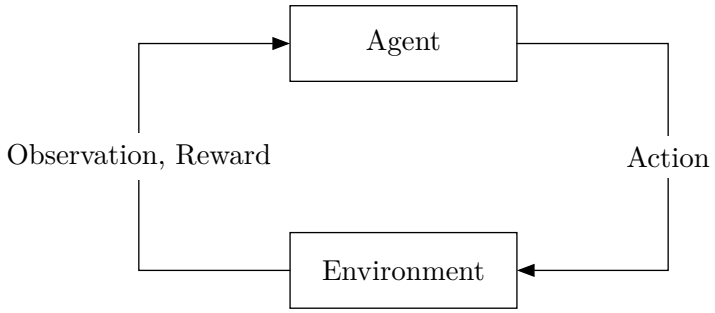


Figure 10: Reinforcement Learning Loop

Unlike supervised learning, where a training data set, i.e., a set of example data labelled by an expert, is used to classify unlabeled data, RL does not require an expert-labelled training data set [35]. Instead, an RL agent learns to predict reward-maximizing actions by exploring the effects of its actions on the environment. In a sense, the agent learns from its own experience. This also implies that, in order to estimate which actions yield a high reward, the agent needs to explore various actions, including those that may make it fail at its task. This issue is commonly referred to as the "trade-off between exploration and exploitation" [35], where exploration means trying potentially suboptimal actions, and exploitation refers to applying actions that are expected to yield a high reward. As there is no universal solution to this trade-off, many RL algorithms, i.e., RL agent implementations, provide parameters ("hyperparemeters") to make the exploration behavior configurable for individual tasks [35], [36].

An important consideration for the application of RL is the choice of environment. Depending on the environment's domain, interactions with a real-world environment may take considerable amounts of time, or even cause damage in case of improper actions. An alternative to training RL agents on real-world environments is the use of a computer simulation. This approach can both reduce the wall-clock time required for training (if the simulation operates faster than the real system), and it can avoid causing damage to a real-world environment. Moreover, computer simulation allows to prototype RL solutions when a real environment is not easily available – as is the case with the cellular handover environment in this thesis.

While a deep understanding of RL algorithms can be helpful, it is not a necessary requirement for applying state-of-the-art RL algorithms. Therefore, this section does not provide a detailed introduction to specific algorithms. However, a few more definitions are relevant for the practical application of RL in this thesis, which will be presented in the following.

- **Policy**: The RL agent's behavior is defined by a policy, representing the decision making that was learned while interacting with the environment [35]. Applying the agent's policy to an observation yields the agent's next action.
- **Value Function**: A value function is learned internally by the RL agent to estimate how desirable a given environment state is over the long term [35]. For instance, taking a specific action in one step might result in a high one-time reward, yet leave the system in a state that makes high future rewards unlikely. A value function prevents this by enabling an RL agent to "plan ahead" and favor actions that promise higher rewards in the long run [35].
- **Reward Function**: The reward function maps the environment's state at a given time to a numeric reward signal for the RL agent. It plays a central role in guiding the agent to optimally solve the task. Its design is crucial for the performance of an RL agent and involves knowledge about the agent's goal [37]. Choosing an inappropriate reward function can make an agent learn a policy that does not solve the given task, despite maximizing the reward signal. This behavior is also known as "reward hacking" [38].
- **Episodic and Continuing Tasks**: Many tasks – such as playing a game – naturally end after a certain event. These tasks are referred to as *episodic* tasks, as they divide into distinct, independent episodes (e.g., multiple rounds of a game) [35]. By contrast, *continuing* tasks may continue infinitely [35]. Controlling handovers in a RAN does not exhibit a natural end and is therefore a continuing task. For practical reasons, when

training RL agents on continuing tasks, the concept of episodes is typically emulated by truncating the task after a set amount of steps [35].

A key advantage of applying RL to the handover optimization problem is that, given a target metric – which is UE QoE in this thesis – the optimization problem for this metric does not have to be solved explicitly, as required for static, domain-specific algorithms. Instead, an RL algorithm dynamically determines a policy to optimize the provided metric. In a complex environment, this policy may exploit environment dynamics that might otherwise remain unnoticed in the process of designing domain-specific algorithms.

2 Related Work

2.1 Handover Optimization

A considerable amount of research has been conducted on optimizing cellular handovers [25], [39]. A relevant selection of these works is presented in this section. The approaches can be grouped into two categories: Those using conventional algorithms (Section 2.1.1) and those based on RL (Section 2.1.2). All presented approaches are summarized and compared in Table 3 and Table 4 at the end of each section, respectively.

2.1.1 Conventional Approaches

Deng *et al.* [14] have conducted UE data rate measurements with several US mobile operators, finding that among all the cells available to a UE at a given time, the selected cell(s) oftentimes did not provide the best data rate in comparison with other available cells. Specifically, at 79% of measurement locations, data rate gains of at least 10 Mbps could be achieved by selecting alternative cells. To prototypically improve the cell selection and increase UE data rates, Deng *et al.* develop a UE-local software approach which they call "Device-Assisted Cell Selection" [14]. It is highlighted that, following the 3GPP-defined handover procedures, the final handover decision is still made by the RAN. The precise approach taken for "assisting" the RAN with handover decisions is not detailed, however, it is likely that UE capability settings are utilized to dynamically exclude frequency bands [5], [40]. To decide between available frequency bands, a profiling mechanism is applied that tracks past data rates for each frequency band and uses this information to predict the frequency band promising the best data rate. In their measurements, Deng *et al.* find that their approach significantly improves data rates, despite occasionally interrupting connectivity for several seconds, thus reducing QoE.

Motivated by the findings of Deng *et al.* [14], Coronado *et al.* [19] implement a bandwidth-aware handover algorithm that can be deployed by the RAN to improve data rates similarly to Deng *et al.*, but for all UEs, without requiring additional UE applications or degrading UE connectivity. Coronado *et al.* base their implementation on the O-RAN architecture[4] which provides open interfaces for softwarized RANs, improving interoperability and speeding up the evolution of RANs. The O-RAN architecture features a "RAN Intelligent Controller (RIC)" which handles real-time network decisions based on a multi-BS central network view [19]. The resource allocation decisions of the RIC, in turn, can be customized by deploying modular applications called *xApps*. Coronado *et al.* utilize this

[4] https://www.o-ran.org/

solution by implementing an O-RAN xApp that optimizes handover decisions for UE data rates, based on the measurement reports the RAN receives from the UEs. They evaluate their solution in a testbed setup and note improved data rates and network fairness in comparison to an off-the-shelf implementation.

Similar to Deng *et al.* [14], Hassan *et al.* [13] have conducted extensive data rate measurements, observing a significant data rate gap between the RAN-selected frequency band and other available bands. The measurements have been taken in five countries (US and Europe) over the course of one year, furthermore including a variety of distinct mobility patterns. In a static mobility setting, the median of the measured data rate gap amounts to 34 Mbps, whereas it comes to 64 Mbps at walking speeds. This indicates that the issue of suboptimal handover decisions has not been solved in practice in the past few years and, furthermore, is significantly impacted by UE mobility patterns. Hassan *et al.* consider data rates with respect to QoE and highlight the impact of taking mobile applications into account to improve QoE via handover decisions. They sketch an O-RAN-based solution, but only implement a UE-local proof of concept similar to the approach of Deng *et al.* [14]. They further bring up the option to incorporate other metrics, such as energy efficiency, into handover decisions – which is approached in this thesis.

Table 3: Handover Optimization Approaches using Conventional Algorithms

Work	Objective	Deployment Strategy
"iCellSpeed" [14]	Optimize data rates for individual UEs	UE-local: Mobile App
"Roadrunner" [19]	Optimize data rates for all UEs	RAN: O-RAN xApp
Hassan *et al.* [13]	Optimize UE QoE (via data rates)	RAN (suggested), UE-local (PoC)

2.1.2 RL-Based Approaches

RL-based handover solutions have been successfully demonstrated for various optimization aspects in different mobile network setups [8], [20], [21], [25]. This section presents three relevant approaches, the concepts of which can be built upon in this thesis. The core design choices of each approach are summarized in Table 4.

Yajnanarayana *et al.* [20] present an RL algorithm for the optimization of 5G handovers using Q-learning [35]. Their motivation is to improve handover decisions with respect to the achievable data rate, which they quantify via the UE's received link beam Signal-to-Noise Ratio (SNR). Unlike the RSRP measurements typically used in handover decisions, the link beam SNR is more directly related to PDSCH data rates [20]. This is because RSRP values are determined using the BS's SSB bursts. Those are primarily intended for cell discovery and beamforming adaption and do not directly mirror the

respective link beam SNR [20]. Unlike RSRPs, however, the effective link beam SNR can only be reported when a UE is already connected to a cell. Thus, it is not directly available to be used in conventional handover algorithms. Yajnanarayana *et al.* apply RL to bridge this gap. They train a central handover agent that is rewarded by a UE's link beam SNR after each handover, thus optimizing this metric via handover decisions. The observation space of the agent consists of the RSRPs reported by each UE. Yajnanarayana *et al.* apply the RL algorithm in simulations and observe an improved link beam gain compared to a conventional RSRP-driven handover algorithm. The precise system simulation software components are not mentioned.

While Yajnanarayana *et al.* optimize data rates in general 5G systems, Mollel *et al.* [21] specifically consider millimeter wave (FR2) deployments with a high cell density. Given the relatively small range of high-frequency cells, frequent handovers may negatively impact UE data rates by temporarily interrupting payload data transfers. Therefore, Mollel *et al.* focus on reducing the number of handovers while maximizing data rates. They apply deep Q-learning [41] with a data-rate-based reward function that penalizes handovers by temporarily reducing the reward signal in each time step in which a handover action has been taken by the RL agent. The agent's observation space includes the SNRs measured by each UE for each cell, as well as the one-hot-encoded serving cell ID of each UE. An evaluation using a commercial radio wave propagation simulator[5] results in increased data rates and 70% less handover procedures compared to a conventional rate-based handover algorithm.

In contrast to the previously presented solutions, Schneider *et al.* [8] do not explicitly focus on handovers. Instead, they consider the cell selection in a coordinated multi-point scenario where multiple BSs can cooperate to improve PDSCH coverage ("Multiple Transmit/Receive Point Operation") [42]. This cell selection problem is a superset of the handover optimization problem, involving both handovers and coordination between BSs. It extends the handover problem by allowing the simultaneous selection of multiple cells from different BSs. Schneider *et al.* apply a Proximal Policy Optimization (PPO) algorithm [36] to a simplistic Python-based system-level simulation, explicitly optimizing QoE, which they model as a logarithmic function of the UE data rate. The reward function therefore is the average QoE of all UEs. Similar to Mollel *et al.* [21], the observation space involves the UE-measured SNRs, as well as – in analogy to the one-hot-encoded serving cell used by Mollel *et al.* – a binary connection vector for each UE, indicating which cells the UE is connected to. Furthermore, the data-rate-based QoE value of each UE is also

[5]Remcom Wireless InSite

included in the observation space. Compared to other heuristic-based coordinated multi-point algorithms, the presented solution achieves a significant improvement of UE QoE. Notably, Schneider *et al.* have made their simulation environment available as open-source software under the MIT license and published it in a conference paper [43]. This makes their environment a useful starting point for the development of a customized, UE-power-aware simulation environment in this thesis.

Table 4: Handover Optimization Approaches using RL

Work	Optimization Metric	RL Algorithm	Observation Space
Yajnanarayana *et al.* [20]	Data rate (link beam SNR)	Q-Learning [35]	RSRP for each UE-cell combination
Mollel *et al.* [21]	Data rate (max.), Handovers (min.)	Deep Q-Network [41]	For each UE: SNRs for all cells, serving cell ID
Schneider *et al.* [8]	UE QoE (based on data rates)	PPO [36]	For each UE: SNRs for all cells, serving cells, QoE

2.2 UE Power Modeling

There are multiple approaches to modeling UE power consumption, involving both empirical and theoretical models. While empirical models are inferred from UE power measurements, theoretical models base their predictions on knowledge about the involved electrical components and their operation. Furthermore, the temporal granularity of a power model may vary depending on the model's use case. In particular, slot-level models predict the average power consumption of a UE in a single time slot, depending on the UE's activity in that slot, while system-level models offer averaged power predictions for a longer period of time, based on higher-level parameters like data rates, traffic patterns, and network configuration. The existing work in this field narrows down to one empirical system-level model for LTE UEs (presented in Section 2.2.1), and a slot-level model for 5G NR UEs (detailed in Section 2.2.2).

2.2.1 LTE System-Level Power Model

Jensen *et al.* [6] present an empirical system-level LTE power model which they obtain via linear regression using power measurements of an "LTE USB dongle" [6]. The UE has been chosen specifically to measure an LTE radio chip in isolation, as opposed to a mobile phone where the radio chip only accounts for a fraction of the overall power consumption [6]. It

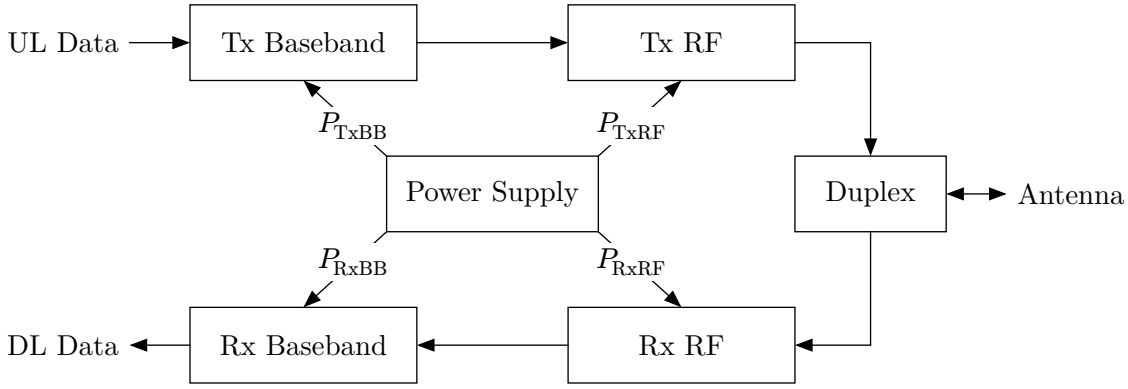


Figure 11: LTE UE Physical Layer Hardware Model (According to and Based on [6])

is noted that the power consumption of the entire device was measured, which, in addition to the radio chip, also includes a USB driver.

The mathematical model definition is based on a simplistic model of the LTE physical layer's hardware, illustrated in Figure 11. The total UE power is represented as the sum of individual component powers, which each depend on the model's input parameters. The core components are baseband and Radio Frequency (RF) circuits for the Transmit (Tx) and Receive (Rx) functionality, respectively. The baseband components primarily handle encoding and decoding of UL and DL data, whereas the RF components process radio signals [6]. The input parameters of the power model, which are Tx/Rx power levels, DL and UL data rate, and the RRC state (connected or idle) each influence the power values of the individual components. The UE power consumption is measured for variations of the input parameters supplied to an LTE BS emulator, and linear regression is applied to estimate the impact of the input parameters on the respective component powers.

The resulting power model is evaluated by applying distinct input parameter sets, yielding an average prediction error of 2.6% and a maximum error of 6%. Notably, the model does not consider DRX [6]. However, Jensen *et al.* have included DRX measurements in a follow-up work [44] in which they adapt the model with mobile phone power measurements. Furthermore, they compare two different LTE radio chip generations and find that the newer chip consumes less power. They attribute this difference to reduced sleep power and improved sleep state transitioning. Finally, their measurements indicate that the radio system accounts for 30–50% of the overall mobile phone power consumption [44].

2.2.2 5G NR Slot-Level Power Model

In a joint effort to reduce 5G NR UE power consumption, the 3GPP members – including major hardware vendors – have conducted a "Study on UE Power Saving" which was published as a 3GPP specification in 2019 [10]. Within their study, a variety of vendors have

performed simulations for various power saving mechanisms. These simulations rely on a slot-level 5G NR UE power model which the vendors have agreed upon and published with the study results. Therefore, it can be assumed that the model is valid for multiple vendors' UEs. Whether the values were obtained empirically or theoretically (using the vendors' hardware designs) is not detailed in the specification. Its temporal granularity and variety of parameters enable the slot-level model to represent a vast amount of configurations, including various bandwidths, antenna configurations, and FR2 operation.

The power model is defined as a mapping of slot activities to slot-average UE power values, as listed in Table 5. The mapping varies for operation in FR1 and FR2. Because the global majority of 5G deployments operate on FR1 (93% in January 2023 [28]), the values for FR2 are not considered below. All power values are measured on a linear, relative Power Unit (PU) scale. Relative means that, compared to an absolute power unit like milliwatts, the PU scale is shifted s.t. one PU corresponds to the UE's minimum power. The mapping of PU to an absolute scale may depend on the considered UE and is not provided in the specification.

Table 5: 3GPP 5G NR FR1 Slot-Averaged UE Power Values [10], [11]

Power State	Relative Power per Slot [PU]
Deep Sleep	$P_{s_{\text{deep}}} = 1$
	$P_{s_{\text{deep}}}^{\text{transition}} = 450, t_{s_{\text{deep}}}^{\text{transition}} = 40$ slots
Light Sleep	$P_{s_{\text{light}}} = 20$
	$P_{s_{\text{light}}}^{\text{transition}} = 100, t_{s_{\text{light}}}^{\text{transition}} = 12$ slots
Micro Sleep	$P_{s_{\text{micro}}} = 45$
	$P_{s_{\text{micro}}}^{\text{transition}} = 0, t_{s_{\text{micro}}}^{\text{transition}} = 0$ slots
1 SSB	$P_{1\text{SSB}} = 75$
2 SSBs or CSI-RS processing	$P_{2\text{SSB}} = 100$
2 SSBs and CSI-RS processing	$P_{2\text{SSB+CSI-RS}} = 170$
PDCCH and 2 SSBs	$P_{r_{\text{CCH}}+2\text{SSB}} = 170$
PDCCH only	$P_{r_{\text{CCH}}} = 100$
PDSCH only	$P_{r_{\text{SCH}}} = 280$
PDCCH + PDSCH	$P_r = 300$
Long PUCCH or PUSCH	$P_t(0 \text{ dBm}) = 250$
	$P_t(23 \text{ dBm}) = 700$
Short PUCCH (1 symb.) or SRS	$P_{t_{\text{short}}} = 0.3P_t$

The slot-level power values in Table 5 assume TDD, 30 kHz SCS, a 100 MHz BWP, 4 Rx antennas, and 1 Tx antenna. This configuration is representative for European mid-band 5G deployments, except for the maximum channel bandwidth, which amounts to 80 or 90 MHz in several European countries, including Germany [28]. The 3GPP model provides scaling factors to adapt slot power values for other configurations. The most relevant scaling factors are listed in Table 6. If the system configuration differs from the defaults, the P_r and P_t values from Table 5 are multiplied with their corresponding scaling factors.

Table 6: Default Configuration and Power Scaling [10]

Property	Default Value	Scaling
DL BWP Bandwidth	100 MHz	For 10, 20, 40, 80 MHz BWPs: $S_{\text{BWP}}(X \text{ MHz}) = 0.4 + 0.6\frac{X-20}{80}$
UL BWP Bandwidth	100 MHz	No scaling
DL Antennas	4 Rx	For 2 RX: $S_{2\text{Rx}} = 0.7$
UL Antennas	1 Tx	For 2 Tx: $S_{2\text{Tx}}(0 \text{ dBm}) = 1.4$, $S_{2\text{Tx}}(23 \text{ dBm}) = 1.2$

With respect to UL transmissions, the model only provides UE power values for transmission powers of 0 dBm and 23 dBm, omitting the interpolation formula for intermediate Tx power values [10]. To obtain a mapping from Tx power values (measured in dBm) to linear-scale UE power values, the Tx power needs to be converted to a linear scale and subsequently mapped to values between $P_t(0 \text{ dBm})$ and $P_t(23 \text{ dBm})$:

$$P_t(x \text{ dBm}) = P_t(0 \text{ dBm}) + \varepsilon \cdot \underbrace{\left(10^{\frac{x}{10}} - 1\right)}_{\text{Lin. power factor}} \tag{2}$$

By substitution with $x = 23$ dBm and solving for the interpolation constant ε, we obtain

$$\varepsilon = \frac{450}{10^{\frac{23}{10}} - 1} \approx 2.267 \tag{3}$$

Therefore, the UE power in PU, given the Tx power in dBm, can be expressed as

$$P_t(x \text{ dBm}) \approx 250 + 2.267 \cdot \left(10^{\frac{x}{10}} - 1\right) \tag{4}$$

Similarly, it is not mentioned how the antenna scaling factor, $S_{2\text{Tx}}(x \text{ dBm})$, is interpolated. Since the factor likely models physical characteristics of the underlying circuits, a logarithmic scaling can be assumed, in line with the definition of $P_t(x)$:

$$S_{2\text{Tx}}(x \text{ dBm}) \approx 1.4 - 0.001 \cdot \left(10^{\frac{x}{10}} - 1\right) \tag{5}$$

3 Power Modeling

Based on the power models presented in Section 2.2, this chapter covers the design and evaluation of a 5G NR UE power model to predict UE power consumption in the context of handover optimization. Section 3.1 analyzes the requirements for such a model. Subsequently, the model is developed in Section 3.2 and evaluated in Section 3.3.

3.1 Requirements

To be applicable to handover optimization, the UE power model developed in this thesis must consider handover/synchronization procedures, as well as payload traffic in RRC connected mode. By contrast, the power consumed by a UE in RRC idle mode is not relevant for handover optimization as handovers only apply to connected UEs. It can therefore be omitted from the power model. Finally, UEs also consume power for RRM measurements. However, this power heavily depends on the RAN configuration and network deployment while only accounting for a small fraction of UE power consumption [10]. It is therefore omitted to reduce model complexity. In addition to the previous considerations, the following enumeration motivates further model requirements.

(1) **System-Level Granularity**: RL research on handovers is commonly performed using system-level simulations [8], [20], [21], [43]. Because system-level simulations for cellular handovers do not model every component of a system in-depth and especially do not operate on a millisecond scale, a slot-level power model cannot easily be integrated into a system-level simulation. Therefore, the developed power model must be able to predict average UE power consumption on a coarser time scale, i.e., a system-level power model is required.

(2) **Accuracy**: The model should accurately represent the effects of its input parameters on UE power consumption to enable realistic predictions.

(3) **Configurability**: In order for the power model to prove useful in handover-driven UE power optimization, it must adequately represent RAN and cell configurations that impact UE power consumption. This requires the model to be configurable with those parameters. However, if there are multiple options for deployment parameters which effect power consumption and one option is predominantly found in today's deployments, that option shall be assumed to reduce model complexity.

(4) **Efficiency**: The model should minimize time- and resource-intensive computations, s.t. it can be embedded in a system-level simulation without significantly impacting simulation performance.

(5) **Extensibility**: The model should be extensible to account for further parameters and mechanisms in future research.

3.2 Design of a System-Level Power Model for 5G NR UEs

As motivated in the previous section, the power model to be developed is required to operate on a system-level granularity. Section 2.2.1 has presented an LTE power model that fulfills this requirement. Yet, without further work, this model is not applicable to 5G NR, which features larger, dynamic cell bandwidths, as well as a variety of power-impacting configuration options. Nevertheless, the linear-regression-based framework that was used to obtain the presented model may be adapted and applied using practical 5G NR UE power measurements. An issue with this approach, however, is that it would require access to a 5G NR BS emulator (in analogy to the LTE BS emulator used in [6]) to practically observe UE powers for the parameter combinations that are necessary to meet the configurability requirement. Such a BS emulator is not easily available and out of scope for this thesis. While commercial 5G NR deployments are available, their configured parameters are fixed by the network operators and do not exhibit sufficient variability to fit an empirical power model [28]. Moreover, with respect to the extensibility requirement, an empirically obtained power model can only be extended via further power measurements. This reduces the scope of feasible model extensions to those mechanisms that are readily available in commercial 5G deployments. Based on these limitations, obtaining an empirical power model is impractical in this work.

An alternative approach is the development of a theoretical system-level model based on 3GPP's slot-level NR power model from Section 2.2.2. This can be achieved by considering the UE activities resulting from a given set of system-level parameters and averaging their power values to predict the expected UE power consumption for the given parameters. This way, the obtained system-level model is highly configurable and remains extensible. The accuracy then depends on the conformity of the deployment parameters with the model's assumptions. If necessary, the model can also be fine-tuned to better represent a particular deployment based on UE power measurements. The derivation of a theoretical system-level power model from the 3GPP model can therefore fulfill the requirements from Section 3.1.

The following two sections cover the design of the desired power model. Specifically, the synchronization and connection establishment procedure involved in handovers is handled in Section 3.2.1, while payload traffic in RRC connected mode is addressed in Section 3.2.2.

As previously mentioned, the prevalent 5G deployment mode is TDD on FR1 [28], which will hence be presumed in both sections.

3.2.1 Synchronization and Connection Establishment

During a handover, the UE synchronizes with and connects to a new cell, temporarily interrupting any payload data transmissions [45]. Therefore, in the context of handover optimization, two metrics are relevant: The time required for the handover, and the average UE power during the handover. Both metrics can be derived by tracing the UE's activity during the synchronization and connection establishment procedure (as outlined in Section 1.1.3 and illustrated in Figure 6). The time T_{Handover} can then be computed as the number of slots occupied by the procedure, and the average UE power consumption P_{Handover} (in PU) as the mean of all slot-average powers. In the following paragraphs, the duration (number of slots) and corresponding sum of slot-average powers is derived for each step of the synchronization and connection establishment procedure. The results are listed in Table 7.

Initial Synchronization: During initial synchronization, the UE repeatedly attempts to receive an SSB (modeled by P_{1SSB}) to synchronize with the cell and decode the MIB. The 3GPP power saving study suggests assuming one SSB/PBCH transmission every 20 ms for power consumption evaluations [10]. Each SSB/PBCH block is also repeated a specific number of times, depending on the carrier frequency and whether the spectrum is shared with another radio access technology [2]. The power model assumes 2 SSBs per slot [10]. Repeated SSBs are transmitted via distinct beams from the BS, enabling the UE to choose the strongest beam [4]. Thus, despite the SSB/PBCH block being repeated several times, the UE might only be able to synchronize once per SSB periodicity. With a periodicity of 20 ms, for 30 kHz SCS, the maximum synchronization time for a UE therefore amounts to 40 slots. Assuming the target cell is not synchronized with the source cell (i.e., the UE attempts synchronization at a random time between two SSB/PBCH transmissions in the target cell), the mean synchronization time is 20 slots. At the end of this step, the UE has synchronized with the cell and decoded the MIB.

Receiving SIB1: The UE monitors the PDCCH for a DCI announcing SIB1 on the PDCCH [2], [5] and at the same time compares the RSRPs of the available beams (determined via the beams' SSBs) to decide on the best beam [4]. Thus, the average power for these slots is modeled by $P_{r_{\text{CCH}}+\text{2SSB}}$. The periodicity for SIB1 transmissions defaults to 20 ms [5], therefore receiving the SIB can be assumed to take at most 40 slots. Again, an average duration of 20 slots will be assumed (19 slots PDCCH/SSBs and 1 slot PDSCH).

Random Access[6]: Depending on the PRACH configuration and assuming a short PRACH format, a PRACH preamble occupies between 2 and 12 symbols [9]. The power model does not specify dedicated PRACH power values, therefore the slot-average power P_t of a long PUCCH transmission is used in this work. The number of PRACH occasions per SSB is decided by the operator and communicated via SIB1 [5]. PRACH resources are aligned with SSBs [9]. For brevity, the average waiting time to receive an SSB (20 slots) will be assumed as the duration between the reception of SIB1 and the configured PRACH occasion. The UE may spend the waiting time before the PRACH occasion in the micro sleep state ($P_{s_{\text{Micro}}}$). Following the UE's PRACH preamble transmission, the BS announces a random access response via the PDCCH and sends it on the PDSCH [2]. The random access response is likely delayed by several slots due to the processing time and scheduling delay at the BS. It will be assumed that the BS requires four slots (2 ms) from receiving the PRACH preamble to responding to it. During these slots, the UE monitors the PDCCH, waiting for the announcement of the response and its subsequent transmission on the PDSCH.

RRC Setup: The random access response includes an upload grant for the RRC connection request [2]. Assuming the default PUSCH time domain resource allocation table[7] is used, the PUSCH slot is scheduled at least four slots and at most seven slots after the reception of the random access response [31]. A scheduling time of four slots will be assumed in this work, i.e., the UE may spend three slots in the micro sleep state before transmitting the connection request on the PUSCH. Following the reception of the connection request at the BS, the RAN sets up the RRC connection. This involves RAN-internal control plane communication and may take a comparably long time. The shortest configurable UE timeout duration for this is 100 ms [5]. Without detailed timing measurements, only an educated guess can be made about the precise network timing. Therefore, the RAN-internal connection setup will be assumed to take 20 ms (40 slots), i.e., one fifths of the shortest configurable timeout duration. During this time, the UE monitors the PDCCH waiting for the announcement of the BS's RRC setup message transmission [5]. Subsequently, the setup message is transmitted on the PDSCH.

RRC Setup Complete: To acknowledge the RRC connection setup, the UE needs to transmit an RRC Setup Complete message [5]. The transmission is initiated by the UE via a SR [2], which can be modelled by $P_{t_{\text{short}}}$. Finally, the UE receives the UL grant via

[6]Random access occasions in the target cell can be configured individually for each UE via RRC signalling before performing the handover, hence a contention-free random access procedure is assumed [5].

[7]"Default PUSCH Time Domain Resource Allocation A for Normal Cyclic Prefix (CP)" [31]

the PDCCH, waits for its PUSCH resource allocation (again assuming a scheduling time of 4 slots), and transmits the RRC Setup Complete message on the PUSCH.

Table 7: Sync. and Conn. Establishment Slot Power Summation

Step	Duration [slots]	UE Energy [PU · slots]
Initial Synchronization	20	$20P_{\text{1SSB}}$
SIB1 Reception	20	$19P_{r_{\text{CCH+2SSB}}} + P_{r_{\text{SCH}}}$
Random Access	26	$20P_{s_{\text{micro}}} + P_t + 4P_{r_{\text{CCH}}} + P_r$
RRC Setup	45	$3P_{s_{\text{micro}}} + P_t + 40P_{r_{\text{CCH}}} + P_r$
RRC Setup Complete	6	$P_{t_{\text{short}}} + P_{r_{\text{CCH}}} + 3P_{s_{\text{micro}}} + P_t$

Summing up the individual slot durations from Table 7 yields $T_{\text{Handover}} = 117$ slots, which translates to 58.5 ms at 30 kHz SCS. This is consistent with a blog post published by the telecommunication company Ericsson, mentioning that handovers can interrupt connections for up to 90 ms [45].

When deriving the average handover power P_{Handover}, one relevant consideration is BWP switching. By default, the 3GPP power model considers a 100 MHz DL BWP. Yet, as illustrated in Figure 9, the BWP used for the connection establishment is substantially smaller [3]. Initially, SIB1 is received using only 24 RBs [3], i.e., 8.64 MHz bandwidth at 30 kHz SCS. A comparably small initial BWP can be assumed for random access and RRC connection establishment. The smallest BWP considered by the 3GPP power model is 10 MHz. Therefore, DL power values need to be multiplied with a 10 MHz BWP scaling factor. One exception in the 3GPP power study is that the minimum scaled value for $P_{r_{\text{CCH}}}$ is fixed at 50 PU, regardless of whether the scaling would result in a lower power value. This applies when using a 10 MHz BWP, allowing to substitute $P_{r_{\text{CCH}}}$ for 50 PU. The sum of all slot-average handover powers can thus be derived as

$$\begin{aligned} E_{\text{Handover}} &= E_{\text{Initial Sync.}} + E_{\text{SIB1}} + E_{\text{RA}} + E_{\text{RRC Seup}} + E_{\text{RRC Setup Complete}} \\ &= 20P_{\text{1SSB}} + 45P_{r_{\text{CCH}}} + 26P_{s_{\text{micro}}} + 3.3P_t + \\ &\quad S_{\text{BWP}}(10\text{ MHz})\left(2P_r + 19P_{r_{\text{CCH+2SSB}}} + P_{r_{\text{SCH}}}\right) \\ &= 3.3P_t + 6255.75 \end{aligned} \tag{6}$$

Based on this, Equation 7 defines the average UE power P_{Handover}, given the UE's transmit power x.

$$P_{\text{Handover}}(x\text{ dBm}) = \frac{E_{\text{Handover}}}{T_{\text{Handover}}} = \frac{3.3P_t(x) + 6255.75}{117} \approx 0.064 \cdot 10^{\frac{x}{10}} + 60.46 \tag{7}$$

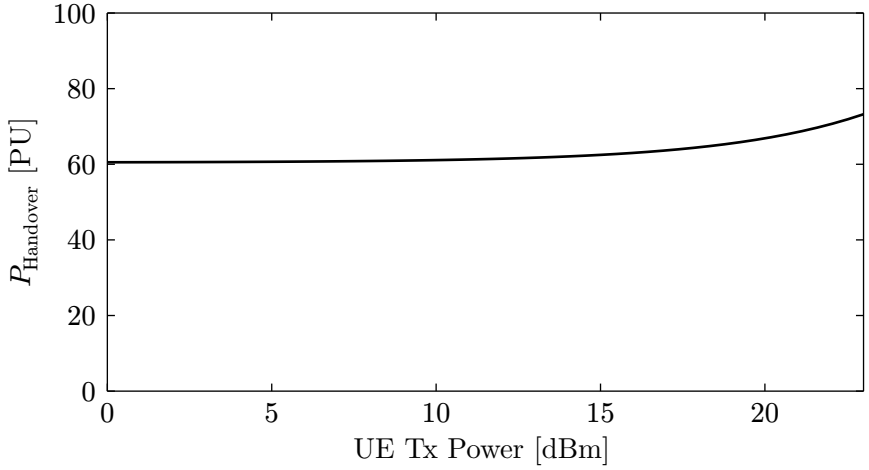


Figure 12: Avg. Handover Power by UE Transmit Power

Figure 12 plots the predicted average handover power over the UE transmit power range. It can be observed that a large portion of the handover power is constant, whereas the transmit power only has a notable impact in the high power range. To obtain a deeper insight into the handover power consumption of a UE, the slot-average powers are plotted over time in Figure 13, assuming a transmit power of 0 dBm. The relatively small transmit time explains the low impact of transmit power on the average handover power. It also becomes evident that a considerable amount of time and UE energy is consumed by waiting for external events, especially in the RRC setup phase. Since the UE spends a relatively large fraction of the handover time in low power states compared to shared channel activity, performing a handover likely consumes less UE energy than actively transmitting and receiving payload data in the same amount of time.

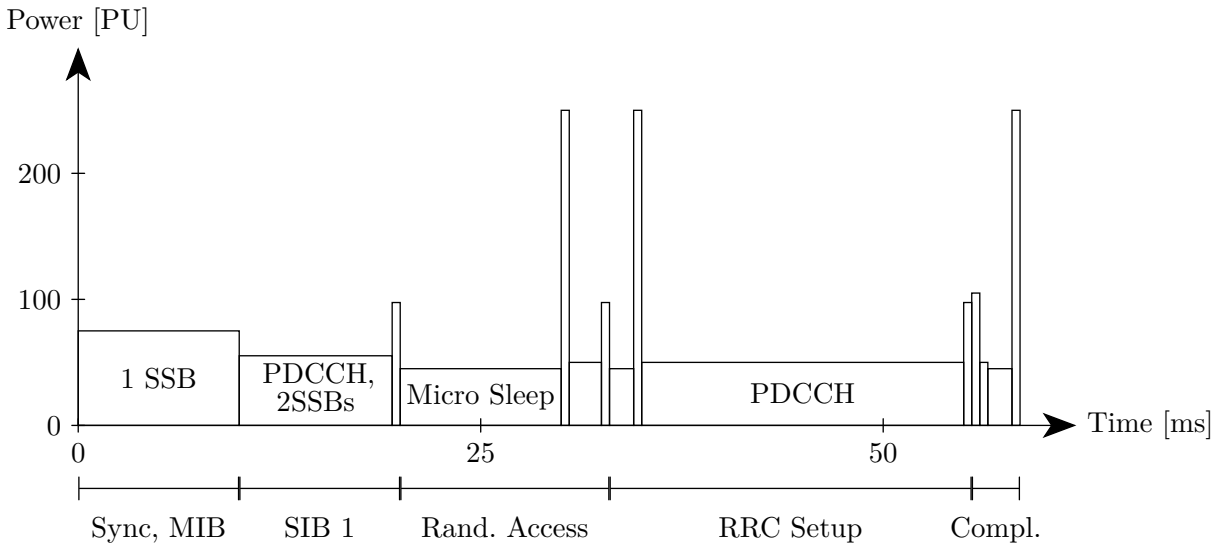


Figure 13: UE Power Consumption for Sync. and RRC Connection Establishment

3.2.2 Payload Traffic

The UE power consumed by payload traffic depends on the UE's transmit power, data rate, and the payload traffic pattern, i.e., how packet arrivals are distributed over time. Moreover, a system-level power model must also incorporate the parameters of power saving mechanisms like DRX and BWP switching. In the following paragraphs, the assumptions made for model development are motivated and the set of model parameters is derived.

The 3GPP power study suggests assuming Poisson-distributed packet arrivals of fixed-size packets for power evaluation[8]. This is not representative for many specific traffic patterns [33], especially considering bursty traffic [47]. However, Poisson-distributed packet arrivals represent an acceptable trade-off between model complexity and accuracy. With respect to DRX, they can be considered a conservative assumption, as their complete randomness likely makes DRX less effective than with bursty traffic patterns. For the scheduling decisions of the BS, the simplifying assumption is made that traffic is scheduled dynamically on a slot-by-slot basis, which is consistent with the randomness of the considered Poisson-distributed packet arrivals. Furthermore, same-slot scheduling is assumed, i.e., a DCI transmission and its corresponding PDSCH allocation occur in the same slot. Regarding HARQ operation, it is presumed that there is enough UL traffic to piggyback DL acknowledgements without distinct PUCCH transmissions. This is realistic when higher-layer protocols with an acknowledgement mechanism are applied. Specifically, in addition to potential transport-layer acknowledgements, the Radio Link Control (RLC) protocol [48] can also be configured to include a link-layer acknowledgement mechanism.

Based on the aforementioned assumptions, the input parameters for the power model are defined in Table 8. Given the large bandwidth available in 5G NR, it is assumed

Table 8: Traffic Power Model Input Parameters

Parameter	Symbol	Unit
Mean DL packet inter-arrival time	T_{P_D}	ms
Mean UL packet inter-arrival time	T_{P_U}	ms
Average Tx power	x	dBm
DRX cycle length	T_C	ms
DRX inactivity timer duration	T_I	ms
DRX On duration	T_{On}	ms
DL BWP size[9]	B_D	MHz

[8] 3GPP calls this traffic model "FTP 3" and assumes a packet size of 0.5 MB [10], [46].

[9] According to the 3GPP power study, S_{BWP} is only valid for $B_D \in \{10, 20, 40, 80, 100\}$ [10].

that a single fixed-size packet can always be transferred in one slot. Hence, payload packet size is not included as a parameter. In the following two sections, the average UE power consumption for the given input parameters is derived from the slot-level 3GPP power model. This is approached by estimating the ratio of DRX-induced sleep time in Section 3.2.2.1 and averaging the slot power values according to the UE's activities in Section 3.2.2.2.

3.2.2.1 Sleep Ratio Estimation

The amount of DRX-induced sleep time has a significant impact on the average UE power consumption. Define the expected relative amount of sleep time as

$$r_{\text{sleep}}\left(T_C, T_{\text{On}}, T_I, T_{P_D}, T_{P_U}\right) \in [0, 1] \tag{8}$$

Appendix A details two approaches to analytically derive r_{sleep} as a function of the input parameters, finding that an analytical representation is hard to obtain. For this reason, a Monte-Carlo simulation will be used to approximate r_{sleep} computationally. A prerequisite for this approach is that it must yield sufficiently accurate results and at the same time fulfill the efficiency requirement from Section 3.1. Notably, r_{sleep} only needs to be computed once per traffic pattern and DRX setting, which facilitates an efficient power model implementation. One notable benefit of estimating the sleep ratio computationally is that, unlike with an analytical solution, the traffic pattern is not limited to Poisson-distributed packet arrivals. Instead, the poisson traffic model can effortlessly be replaced with a different traffic model in the Monte-Carlo simulation, in accordance with the extensibility requirement.

The decision for a Monte-Carlo approach requires the choice of a programming language. Because the power model is intended to be applied in a machine learning context, where the predominant programming language is Python, the Monte-Carlo simulation should be compatible with Python. Thus, an initial version of the simulation has been written in Python. However, an equivalent C++ implementation yields an approximate speedup of 180, making C++ a more suitable choice with respect to the efficiency requirement. Additionally, when using C++, the simulation loop can be parallelized effortlessly by applying OpenMP. To maintain Python compatibility, the resulting C++ function has been exposed as a Python function using Cython[10]. Listing 1 details the core simulation logic. The full code is available online in the repository of this thesis[11].

[10] https://cython.org/

[11] https://github.com/bjoluc/5g-handover-optimization

```
class UE {
 private:
  Timer inactivityTimer;
  Timer onDurationTimer;
  PoissonArrivals ulArrivals;
  PoissonArrivals dlArrivals;
  bool isDlPacketBufferedAtBS = false;

 public:
  UE(unsigned int iatUl, unsigned int iatDl, unsigned int onDuration,
     unsigned int inactivityTimer)
      : inactivityTimer(Timer(inactivityTimer)),
        onDurationTimer(Timer(onDuration)),
        ulArrivals(PoissonArrivals(iatUl)),
        dlArrivals(PoissonArrivals(iatDl)) {};

  bool isActive = true;

  void startNewCycle() {
    // On duration starts in this slot
    this->onDurationTimer.set();
    this->isActive = true;

    // UE receives buffered DL traffic, if any
    if (this->isDlPacketBufferedAtBS) {
      this->inactivityTimer.set();
      this->isDlPacketBufferedAtBS = false;
    }
  }

  void step() {
    bool doesOnDurationTimerExpire = this->onDurationTimer.step();
    bool doesInactivityTimerExpire = this->inactivityTimer.step();
    bool doesUlPacketArrive = this->ulArrivals.step();
    bool doesDlPacketArriveAtBS = this->dlArrivals.step();

    // Handle on duration and inactivity timer expiry
    if ((doesOnDurationTimerExpire || doesInactivityTimerExpire) &&
        !(this->inactivityTimer.isRunning() ||
          this->onDurationTimer.isRunning())) {
      this->isActive = false;
    }

    if (doesUlPacketArrive) {
      if (!this->isActive) {
        this->isActive = true;

        // UE receives any buffered DL traffic too:
        this->isDlPacketBufferedAtBS = false;
      }

      this->inactivityTimer.set();
    }

    if (doesDlPacketArriveAtBS) {
      if (this->isActive) {
        this->inactivityTimer.set();
      } else {
        this->isDlPacketBufferedAtBS = true;
      }
    }
  }
};
```

Listing 1: Sleep Ratio Simulation Logic

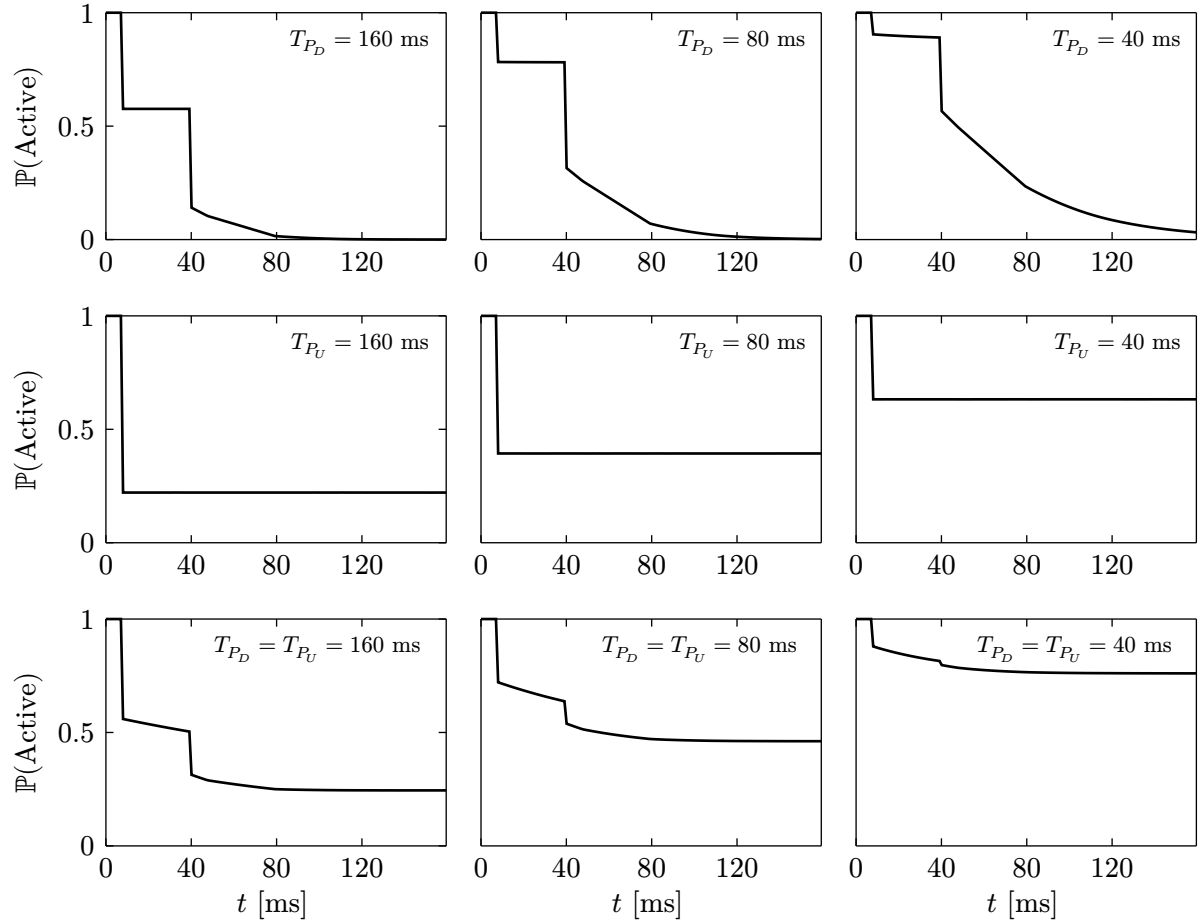


Figure 14: UE Activity Probability for $T_C = 160$ ms, $T_I = 40$ ms, and $T_{\text{On}} = 8$ ms

Applying the obtained sleep ratio simulation model, r_{sleep} can be approximated efficiently by simulating a configurable number of DRX cycles. The simulation model can also be used to gain a deeper understanding of the effects of different UL and DL packet rates on UE activity for a given set of DRX parameters. The 3GPP power study suggests $T_C = 160$ ms, $T_I \in \{40, 100\}$ ms, and $T_{\text{On}} = 8$ ms as a reference DRX configuration for UE power evaluation [10]. In Figure 14, these DRX settings (choosing $T_I = 40$ms) have been applied with varying packet rates and the UE activity probability has been plotted over the DRX cycle. It can be observed that the UL packet rate has a direct impact on UE activity over the entire DRX cycle, as UL traffic wakes up the UE on arrival. By contrast, DL traffic in the absence of UL traffic primarily activates the UE at the beginning of a DRX cycle due to DRX-induced buffering at the BS, resulting in a higher overall sleep ratio and latency. When DL traffic is combined with UL traffic, the amount of sleep time is reduced and less DL packets arrive while the UE is inactive, decreasing the average DRX-induced latency.

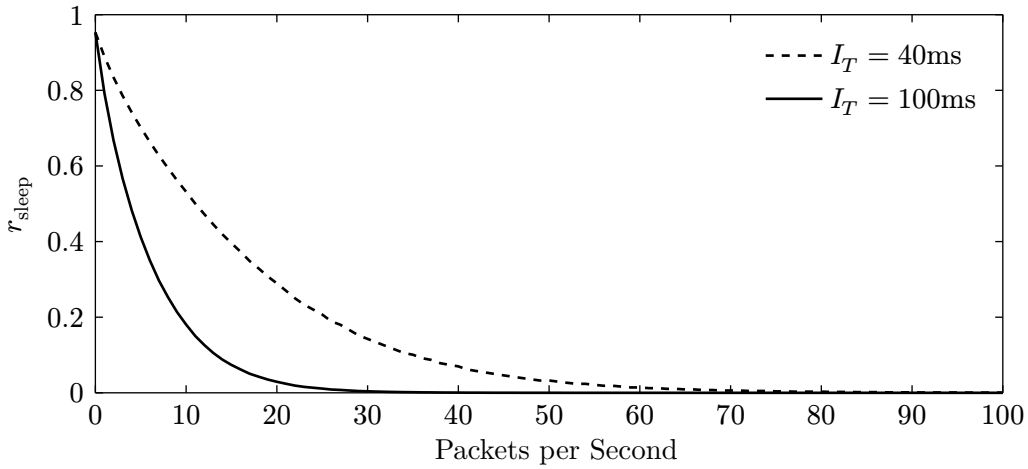


Figure 15: Sleep Ratio over Joint UL/DL Avg. Packet Rate

Due to the communication overhead of higher-layer protocols, UL and DL packet rates are likely to be related for many applications. Hence, when estimating the practical impact of DRX settings, it is natural to consider similar UL and DL packet rates. Figure 15 plots r_{sleep} over the average packet rate (identical for UL and DL), comparing the impact of a short (40 ms) and a long (100 ms) inactivity timer on the UE's sleep ratio. The plot shows that for a given average packet rate, the short inactivity timer expires more frequently than the long inactivity timer, resulting in a higher average sleep ratio. As the average packet rate increases, the sleep ratios for both inactivity timer values approach zero, since the UE is permanently kept awake by traffic.

To obtain highly accurate plots, Figure 14 and Figure 15 have been generated by simulating 10^7 and 10^6 cycles, respectively. Nevertheless, r_{sleep} can be estimated sufficiently accurate (± 0.0023) by only simulating 10^3 cycles for the given DRX settings[12]. On a modern notebook, this takes about 72 ms on a single core and roughly 12 ms in parallel. The time overhead required for the sleep ratio estimation is thus acceptable with respect to the efficiency requirement. Furthermore, r_{sleep} can be cached for each considered traffic pattern and DRX setting to avoid redundant computations on subsequent power estimations.

3.2.2.2 Slot Power Summation and Averaging

In Section 3.2.1, the handover power model was obtained by considering the individual UE activities on a slot basis and summing up their slot-average power values. A similar approach can be used for the traffic model, based on the input parameters from Table 8. This entails summing up the slot powers for an arbitrary time interval T (in ms) and

[12]These values were obtained by running 10^4 estimations for $T_{P_D} = T_{P_U} = 80$ms and computing the std. deviation of r_{sleep}, as well as the mean computation time.

dividing the sum by the number S of slots in T (where $S = 2T$ due to the assumed SCS of 30 kHz). The sum of slots is known to be composed of a set of distinct power states, as detailed in Equation 9 below.[13]

$$S = 2T = S_{\text{sleep}} + S_{\text{PDCCH+PUCCH}} + S_{\text{PUSCH}} + S_{\text{PDCCH}} + S_{\text{PDCCH+PDSCH}} \tag{9}$$

The term $S_{\text{PDCCH+PUCCH}}$ in Equation 9 assumes that every slot contains a PDCCH and therefore in every PUCCH slot, the UE also receives the PDCCH. The 3GPP power model represents this combination as the sum of $P_{r_{\text{CCH}}}$ and $P_{t_{\text{short}}}$ [10]. In the following paragraphs, each component of Equation 9 will be derived individually.

UL Traffic: Per the assumptions, each UL packet requires a short PUCCH for the scheduling request and one PUSCH slot. The Poisson distribution yields the expected number of UL packets in T ms as $\frac{T}{T_{P_U}}$, thus

$$S_{\text{PDCCH+PUCCH}} = S_{\text{PUSCH}} = \frac{T}{T_{P_U}} \tag{10}$$

DL Traffic: Per the assumptions, one PDCCH+PDSCH slot is counted for each DL packet. Based on the Poisson distribution, the expected number of DL packets in T ms is $\frac{T}{T_{P_D}}$, therefore

$$S_{\text{PDCCH+PDSCH}} = \frac{T}{T_{P_D}} \tag{11}$$

The amount of sleep time is given by r_{sleep}, yielding

$$S_{\text{sleep}} = r_{\text{sleep}} \cdot S \tag{12}$$

The remaining active time is spent either in one of the active states counted above, or by receiving PDCCH. The amount of PDCCH-only slots can therefore be expressed as the remaining number of active slots after considering all other activities:

$$S_{\text{PDCCH}} = \left(1 - r_{\text{sleep}}\right) \cdot S - \left(S_{\text{PDCCH+PUCCH}} + S_{\text{PUSCH}} + S_{\text{PDCCH+PDSCH}}\right) \tag{13}$$

Substituting the slot counts in Equation 9 by the derived terms above yields a formula that simplifies to S as expected. Therefore, the individual slot counts properly decompose the total number S of considered slots. In addition to the power consumption in individual slots, there is a power overhead $P_{s_{\text{deep}}}^{\text{transition}}$ for ramp-down and ramp-up procedures to

[13]This decomposition does not consider SRS transmissions and CSI-RS processing to reduce model complexity.

transition into and out of a sleep state. To account for this, the number of sleep transitions in time T needs to be determined. This can be achieved by counting the number of sleep periods in the sleep ratio simulation and normalizing it to a fixed amount of time, e.g., a cycle time. Using this approach, define $n_{\text{sleep}}\left(T_{P_D}, T_{P_U}, T_C, T_{\text{On}}, T_I\right)$ as the average number of transitions into a sleep state within a DRX cycle.

The UE energy consumption in S slots is obtained by summing up the slot-average powers for each slot activity, as well as the additional sleep transition power. Moreover, to respect the DL BWP size B_D, all DL power values need to be scaled with the corresponding BWP scaling factor. Dividing the energy sum by S then yields the average UE power consumption for the given input parameters, as derived in Equation 14.

$$
\begin{aligned}
&P_{\text{traffic}}\left(T_{P_D}, T_{P_U}, T_C, T_{\text{On}}, T_I, x, B_D\right) \\
&= \frac{1}{S}\Bigg(P_{s_{\text{deep}}} \cdot S_{\text{sleep}} + P^{\text{transition}}_{s_{\text{deep}}} \cdot n_{\text{sleep}} \cdot \frac{T}{T_C} + \\
&\quad S_{\text{BWP}}(B_D)\left(P_{r_{\text{CCH}}} \cdot \left(S_{\text{PDCCH}} + S_{\text{PDCCH+PUCCH}}\right) + P_r \cdot S_{\text{PDCCH+PDSCH}}\right) + \\
&\quad P_{t_{\text{short}}}(x) \cdot S_{\text{PDCCH+PUCCH}} + P_t(x) \cdot S_{\text{PUSCH}} \Bigg) \\
&\approx \frac{1.473 \cdot 10^{\frac{x}{10}} - 0.375 B_D + 148.5}{T_{P_U}} + \frac{0.75 B_D + 25}{T_{P_D}} + 0.75 B_D + 25 \\
&\quad - r_{\text{sleep}}\left(T_C, T_{\text{On}}, T_I, T_{P_D}, T_{P_U}\right)(0.75 B_D + 24) \\
&\quad + n_{\text{sleep}}\left(T_C, T_{\text{On}}, T_I, T_{P_D}, T_{P_U}\right)\frac{225}{T_C}
\end{aligned}
\tag{14}
$$

For further use in predictions and simulations – specifically in the RL-based handover optimization in this thesis – the obtained power model has been implemented as a Python module. It includes the simulation-based sleep time estimation from Section 3.2.2.1, handles the caching of r_{sleep} and n_{sleep} estimates, and implements the power formula from Equation 14. It has been applied to compute UE power predictions for a variety of parameters and the results have been plotted over the average packet rate in Figure 16. The considered DRX parameters have been adopted from the plots in the previous section to allow for a direct comparison of the predicted power consumption with the sleep ratio estimates in Figure 15. Three major effects can be observed in Figure 16:

(1) The DRX-induced sleep time drastically decreases the average UE power consumption at low packet rates, while having no notable impact at higher packet rates.

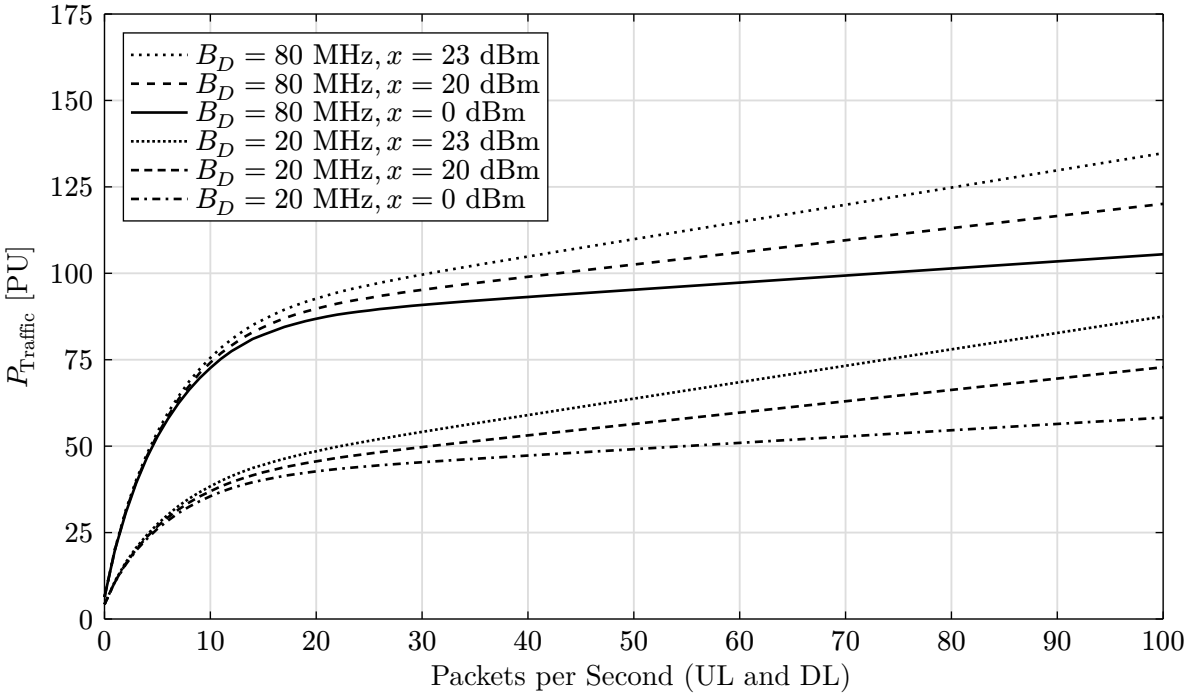


Figure 16: UE Power over Avg. Packet Rate ($T_C = 160\text{ms}, T_{\text{On}} = 8\text{ms}, T_I = 100\text{ms}$)

(2) The size of the configured DL BWP (B_D) significantly impacts the overall power consumption. The impact of the packet rate on this effect is insignificant for the comparably low packet rates considered in Figure 16. Specifically, the power gap between a 20 MHz BWP and an 80 MHz BWP is almost constant for packet rates in the range of 30 to 100 packets per second. This can be explained by the large amount of PDCCH slots compared to only a small, packet-rate-dependent number of PDSCH slots.[14]

(3) The effect of a UE's transmit power on its power consumption increases linearly with the average UL packet rate and is independent of the DL BWP size. It is exponentially impacted by the given dBm transmit power value (due to the logarithmic nature of the dBm scale).

3.3 Measurement-Based Validation

As mentioned at the beginning of this chapter, the configuration of commercial 5G deployments is operator-defined and does not exhibit sufficient variability to fit an empirical power model. Yet, the theoretical power model that was designed in the previous section can be compared against UE power measurements obtained in a commercial 5G deployment to evaluate the model's accuracy. For this purpose, the mobile operator Deutsche Telekom AG has volunteered to provide the necessary resources, including SIM

[14] At 100 DL packets per second, on average, there is only one PDSCH slot every 20 slots.

cards that are enabled for 5G SA and 5G NSA, respectively. This way, the power model predictions can be validated for a specific set of network parameters. Furthermore, the model's applicability to 5G NSA, which is the prevalent mode of operation in commercial 5G deployments [28], can be evaluated.

Conducting UE power measurements requires a suitable UE. Because the power consumption of the baseband processor in a commercially available mobile phone cannot easily be measured in isolation, Jensen *et al.* [6] have applied an "LTE USB dongle" [6] for their empirical power model. Such a product is not readily available with 5G support. However, there are external 5G modules with an M.2 form factor which can be controlled via a USB connection. Therefore, the decision has been made to use a Waveshare SIM8202G-M2 5G HAT[15], which consists of a Printed Circuit Board (PCB) providing various external interfaces to a Simcom SIM8202G-M2 5G module. The 5G module, in turn, features a Qualcomm Snapdragon x55 5G baseband processor[16] which is also used in a number of commodity mobile phones, including Samsung's Galaxy S20 and Apple's iPhone 12. Hence, the considered baseband processor is likely to be representative for a broad number of UEs.

In order to enable isolated power measurements of the radio module, the PCB has been modified to separate the module's power supply from the PCB's 4.2 V rail[17] [49]. External sockets have been mounted to separately access the 4.2 V rail and the module's power supply pins. Given the near constant voltage supply of 4.2 V, the module's power consumption can be assumed to be proportional to its current draw. Thus, measuring UE power consumption translates to measuring the UE's electrical current. In an initial attempt, a Siglent SDM 3045X digital multimeter has been used to measure the current draw of the radio module. However, the multimeter's maximum sampling rate of 150 Hz has proved to be too coarse to capture short current peaks and gain sufficient insight into the UE's activities. For this reason, the multimeter has been replaced with a TiePie Handyscope HS3 USB oscilloscope with a maximum sampling rate of up to 100 MHz. To infer the module's current draw from the oscilloscope's voltage measurements, a 0.5 Ohms shunt resistor is applied. Two channels of the oscilloscope are used to record the voltage between ground and both sides of the shunt. Subsequently, a Python-based measurement script is applied to compute currents from the measured voltage differences. To obtain sufficiently detailed measurements, a sampling frequency of 32 kHz is used, resulting in 16 samples per 5G NR slot at 30 kHz SCS. Notably, due to the use of a USB oscilloscope, the measurement setup is portable and can be powered by a notebook.

[15]https://www.waveshare.com/sim8202g-m2-5g-hat.htm
[16]https://www.qualcomm.com/products/technology/modems/snapdragon-x55-5g-modem
[17]denoted as "VBAT" in the schematic [49]

For the experiments, the modem is connected to a notebook via two USB cables for power supply and data transfer, respectively. The data USB connection serves two purposes. First, it is used to transfer data plane traffic. Second, control plane messages can be exchanged with the modem via the Qualcomm Mobile Station Modem Interface (QMI). This allows to configure the modem, control radio connections, and retrieve metadata about the modem's state[18]. For the generation of payload data in compliance with the power model's traffic assumptions, fixed-size random-byte User Datagram Protocol (UDP) packets with exponentially distributed inter-arrival times are exchanged between the UE and an external server. The UDP protocol has been chosen to avoid transport-layer communication overhead.

An initial set of current measurements has been obtained without payload traffic[19] to estimate RAN parameters and analyze the UE's power states. The measurements have been performed with a 5G NR cell on the n78 frequency band (TDD) using a center frequency of 3626.4 MHz. The cell's bandwidth is not exposed via the QMI protocol[20], but the mobile operator is known to support a bandwidth of 90 MHz on the given band [28]. BS and UE were in line of sight and the UE reported an RSRP of −62 dBm.

Figure 17 shows the UE's current in RRC connected mode over a period of 1.28 seconds. Because the signal has been reduced for plotting, very short high-current peaks are not depicted accurately. The current curve indicates a periodic activity, alternating with a low-power state, characterized by a current of 37 mA. It must be assumed that the activity

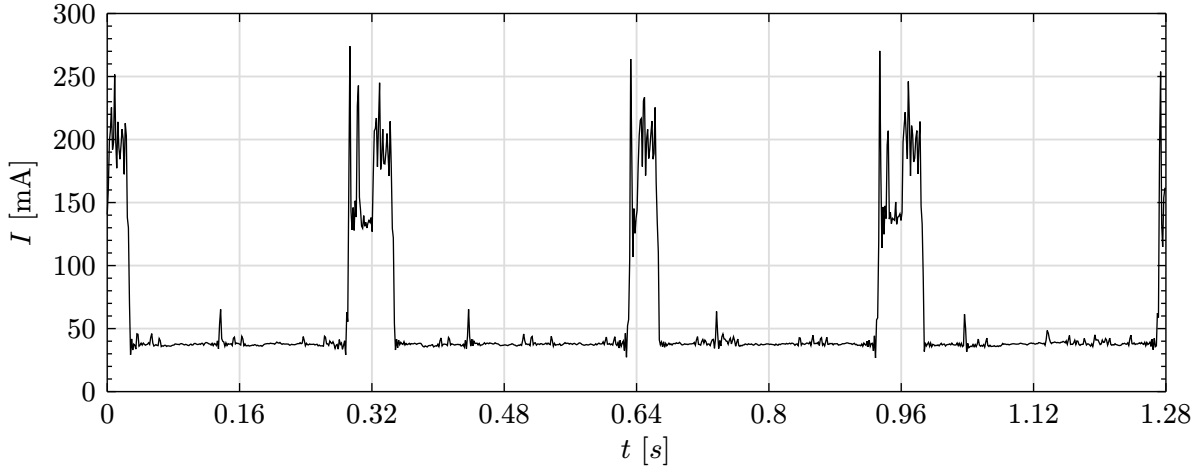


Figure 17: RRC Connected Mode Current without Payload Traffic

[18]This has been implemented using the `libqmi` library (https://www.freedesktop.org/wiki/Software/libqmi/) via its GLib Python bindings.

[19]A DL packet has been received every four seconds to keep the UE in RRC connected mode during the measurements.

[20]The QMI protocol specifies a response field for the channel bandwidth, but the given UE does not return this information for 5G NR cells.

is caused by the DRX On duration and the low-power state represents the UE's deep sleep state. Presuming a DRX cycle duration of $T_C = 320$ ms, it is evident that the UE wakes up a variable time before the beginning of the On duration. This may be required to maintain the UE's synchronization or perform RRM measurements and is not considered by the developed power model. To examine the UE's active power states, the third On duration from Figure 17 is depicted on a finer time scale in Figure 18. Apparently, the DRX On duration starts at 15 ms in the plot and ends a bit after 35 ms, followed by the deep-sleep ramp-down procedure. This indicates that – according to the 3GPP-specified set of configurable timer values – the On duration timer has been configured to $T_{\mathrm{On}} = 20$ ms. The operator has confirmed these DRX settings and furthermore provided the inactivity timer value $T_I = 100$ ms. It has also been noted that a minimum On duration of 10 ms has been configured to the BS which, in turn, adapts this value as needed for channel estimation (eg., SRS or CSI-RS). With respect to channel estimation, the recurring current peaks in Figure 18 (starting at 17 ms) have proven to be SRSs, as a group of SRSs (four in this plot) occurs every 80 slots, which is the operator-confirmed SRS periodicity for TDD cells.

During the On duration, the UE is required to monitor the PDCCH. This can be observed as a recurring current peak at slot frequency, i.e., every 0.5 ms. In PDCCH-only slots, this peak is followed by a micro-sleep period for the rest of the slot duration. Notably, according to the currents in Figure 18, every fifths slot does not seem to be configured for a PDCCH transmission. These are also the slots in which a SRS is transmitted. Further measurements involving UL transmissions – which are not detailed here for brevity – indicate that every fifths slot is configured as UL-only, preceded by four DL-only slots.

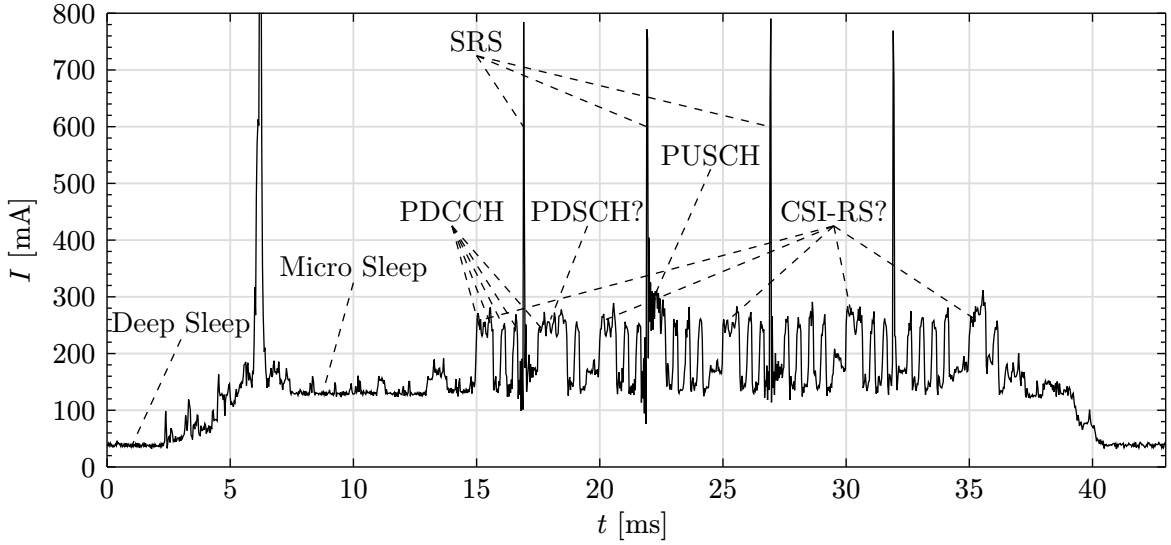


Figure 18: DRX On Duration Current

A number of slots do not involve micro-sleep periods. Particularly, in every 10th slot (starting at 15 ms in Figure 18), the PDCCH reception is followed by another activity at a similar power level. Since the DRX inactivity timer is not triggered by these activities, payload PDSCH transmissions can be ruled out as a reason. A possible cause, however, may be the frequent processing of received CSI-RS signals. Additionally, Figure 18 shows prolonged activity at 17.5 ms and at 22 ms. The activity at 17.5 ms may be caused by a less frequent RRC-related DL event that does not trigger the inactivity timer, like the reception of an SIB or a dedicated RRC message on the PDSCH [32]. By contrast, the activity at 22 ms causes a slightly higher current and is located after the SRS in an UL-only slot. It might therefore be a low-transmit-power UL (PUSCH or long PUCCH) transmission.

In the previous paragraphs, relevant RAN-configured parameters have been derived from fine-grained UE current measurements and the currents have been linked to specific activities at the slot level. In order to evaluate the accuracy of the previous section's power model, a mapping of the model's relative PU values to absolute UE currents is required. This can be established using the slot-average currents of the individual slot activities. For this purpose, Figure 19 averages the currents from Figure 18 per slot to enable a direct comparison with the PU values from the 3GPP power model. Moreover, the plot also marks the slot-average current for each of the observed activities.

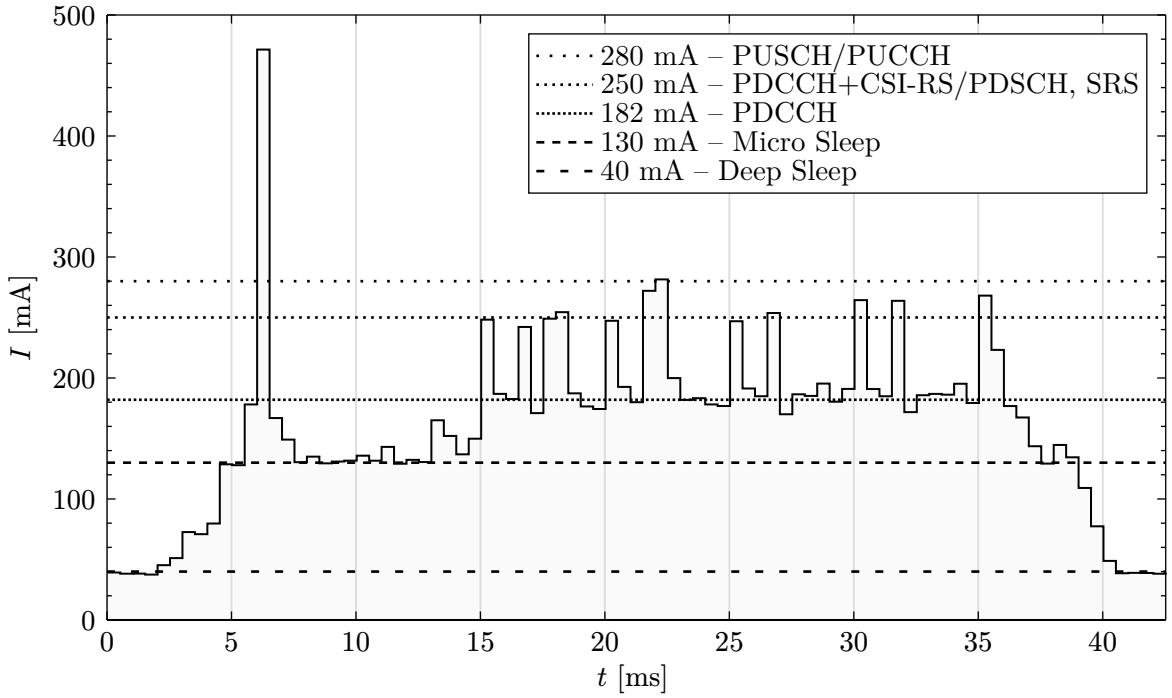


Figure 19: Slot-Averaged DRX On Duration Current

Per the 3GPP power model's definition of a power unit, 1 PU represents the deep sleep power. Hence, the linear mapping of PUs to Milliamperes can be expressed as

$$I(p) = a(p-1) + I_{s_{\text{deep}}} \tag{15}$$

where $I_{s_{\text{deep}}} = 40$ mA is the UE's deep sleep current. The scaling factor a in Equation 15 may be obtained by substitution, using one of the determined slot-average current values and its corresponding PU value. In Figure 19, the most frequently observed non-sleep power state is $P_{r_{\text{CCH}}}$ with an equivalent current of 182 mA. It is therefore used to obtain the value of a in Equation 16 below.

$$I\left(S_{\text{BWP}}(90\text{ MHz})P_{r_{\text{CCH}}}\right) = 182 \;\Rightarrow\; a \approx 1.55 \tag{16}$$

The BWP scaling for 90 MHz is applied to represent the considered cell's bandwidth in the 3GPP power model's PU value. Notably, the 3GPP model does not specify S_{BWP} for 90 MHz. Nevertheless, the linear definition of S_{BWP} for other bandwidths suggests that the scaling is not substantially inaccurate for 90 MHz either.

Finally, the resulting PU-to-mA mapping $I(p)$ can be used to compare predictions of the system-level UE power model with the average measured UE currents for various conditions. Due to the complexity of accurately timing and measuring the synchronization and connection establishment procedure, the handover power model from Section 3.2.1 is not validated in this work. By contrast, the traffic power model from Section 3.2.2 can be validated more easily by exposing the UE to varying traffic patterns and averaging the measured UE current. Because DRX and BWP settings, as well as the Tx power, are RAN-configured, only traffic patterns can be fully controlled in the given experiment setup. Hence, the measurements in this section primarily consider the avg. UL/DL packet rate, in analogy to the power model plots in Figure 16.

As mentioned above, in accordance with the model assumptions, the generated payload traffic consists of Poisson-distributed, fixed-size UDP packets with a configurable inter-arrival time. Regarding the packet size, the 3GPP-defined FTP 3 traffic model assumes the transmission of files with a fixed size of 512 kB. Payloads of this size cannot be transmitted in a single UDP packet. Introducing multiple subsequent UDP packets, however, does not comply with the exponential inter-arrival time assumption and therefore might affect the DRX sleep time. Furthermore, a typical network-enforced Maximum Transmission Unit (MTU) is 1500 Bytes, resulting in fragmentation of larger packets [34]. Hence, to avoid fragmentation, the payload size of a DL UDP packet has been fixed to 1 kB. Furthermore, to mimic typical downlink-heavy traffic, the UL packet size has been set to 0.25 kB. A

preliminary measurement series at 200 packets per second has shown that a varying packet size in the range from 16 Bytes to 1 kB has no noticeable impact on the average UE power consumption.

Due to the oscilloscope's limited buffer size, current measurements at the given sample rate can only last up to four seconds. To determine the UE's average current for a given packet rate, the oscilloscope is therefore used to repeatedly measure a four-second interval, followed by a gap of about 250 ms, respectively, to transfer, process, and average the measurements. For each considered packet rate, a total time of 16 seconds is measured as a batch of four-second measurements. For packet rates below 21 packets per second, this time has been increased to 64 seconds to account for the large current variance stemming from infrequent packet arrivals and DRX. To further reduce variance due to varying cell loads and BS-configured UE transmit powers, the above measurements have been repeated nine times[21] per packet rate, for rates from 1 to 200 packets per second, yielding a total measurement time of 10.4 hours. The order of measured packet rates was further randomized in each iteration to avoid capturing time-related effects in the resulting average power curve. The measurement procedure has been performed once for a 5G SA setup and once for a 5G NSA setup in the same cell (n78, 3626.4 MHz), at the same location (-62 dBm RSRP), and at the same time of day (over night). Notably, the RSRP of the primary LTE cell (band 3, FDD, 1720 MHz UL, 1815 MHz DL, 20 MHz bandwidth) in 5G NSA mode was only reported as −79 dBm.

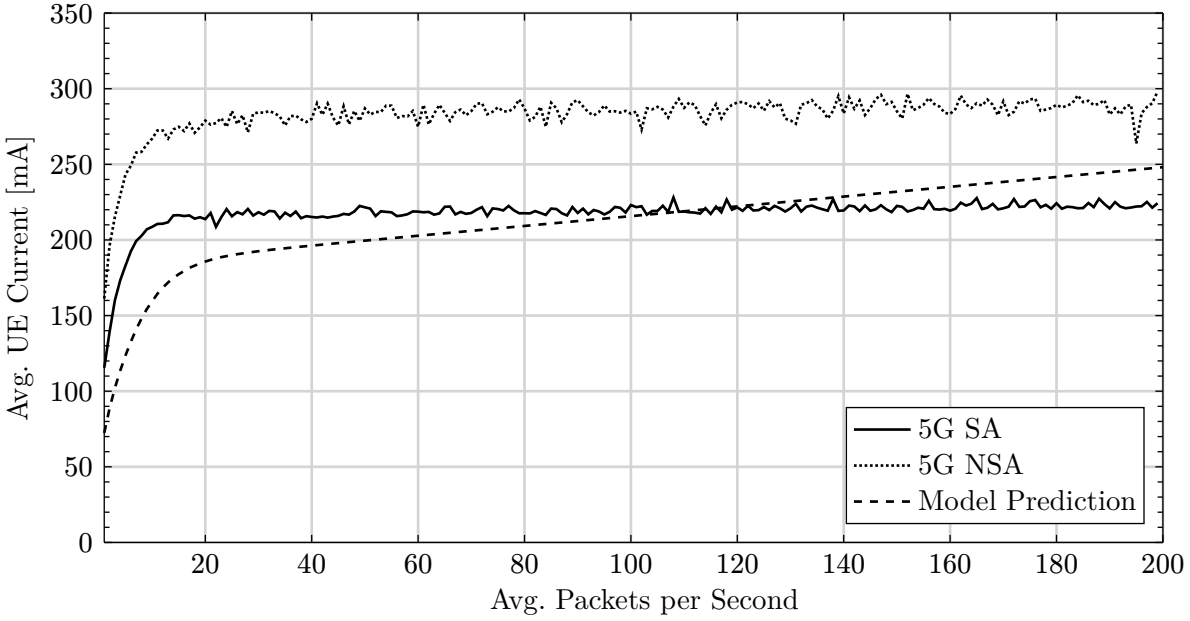


Figure 20: Comparison of Traffic Power Model Predictions with 5G SA/NSA Measurements

[21] This number was chosen due to practical time constraints. Thanks to the team of actiVita Paderborn for offering their gym as a measurement location for several nights!

The resulting current measurements have been plotted in Figure 20, including the corresponding mapped traffic power model predictions $I(P_{\text{traffic}})$ for an assumed transmit power of 0 dBm[22] and a BWP size $B_D = 90$ MHz. Compared to the 5G SA measurements, the power model predictions exhibit a mean error of 7.8% for the given parameters. Four key observations can be made:

(1) While the power model's current predictions have a strong linear dependency on the average packet rate, the measurements only show a minor trend.
(2) The general effect of the DRX mechanism on UE power consumption is adequately represented.
(3) The DRX-based average current reduction at low packet rates is less significant than predicted by the model.
(4) 5G SA operation consistently exhibits lower average currents than 5G NSA operation in the considered setup. The mean difference is 65 mA and the mean 5G SA power reduction compared to 5G NSA is 23%.

Notably, all effects predicted by the model can also be observed in the measurements for both 5G SA and 5G NSA. However, the intensity of these effects does not match the experiment results, as described in observations (1) and (3). With respect to Observation (1), the different effect sizes may partially stem from dynamically configured bandwidth parts, intelligent BS scheduling decisions (e.g., combining multiple successive small packets into one transmission), or the use of an overly small packet size in the measurement setup. Regarding Observation (3), the reduced effect size of DRX-induced sleep time on the UE power consumption is likely related to the UE's activities before the DRX On duration (as observed in Figure 17 and Figure 18). It is also impacted by active-mode activities such as CSI-RS processing and SRS transmissions, which are not considered by the power model. Furthermore, the ratios between the 3GPP-defined power values for various slot activities are not guaranteed to equally apply to every UE, introducing further inaccuracy and possibly further impacting the effect sizes in observations (1) and (3). With regard to Observation (4), the reduced power consumption of 5G SA compared to 5G NSA is likely caused by additional UE activities in 5G NSA mode, introduced by the carrier aggregation of an LTE frequency band with a 5G NR band.

Another power-relevant model parameter worth validating is the UE's transmit power x. Unlike packet rates, however, the transmit power cannot be chosen arbitrarily by the UE as it is ultimately configured by the BS [32]. Yet, the path loss can be increased

[22]The UE does not expose its transmit power for 5G NR cells via its QMI interface. The assumption of 0 dBm is motivated by the high RSRP values resulting from the line-of-sight communication with the BS.

incrementally by stepwise moving the UE away from the BS, leading the BS to configure a higher UE transmit power. Figure 21 plots the results of such an experiment[23] with a fixed packet rate of 100 packets per second (for UL and DL). At that rate, under the Poisson assumption for packet arrivals, it is highly unlikely to trigger the DRX inactivity timer of 100 ms. Hence, the Poisson assumption has no notable impact on UE power consumption and the packet inter-arrival time has been fixed to 10 ms to reduce traffic-based UE power variance in the measurements. The experiment has been conducted by repeatedly measuring four-second current averages, each followed by a short interval to move the UE further away from the BS, until the UE was handed over to another cell by the RAN. As mentioned before, the configured transmit power is not exposed by the particular UE and therefore has not been captured. Instead, the measurements have been plotted over the UE-determined RSRP. Furthermore, the current-mapped power model predictions at the given traffic rate have been plotted for the model's minimum and maximum transmit power. The maximum transmit power of 23 dBm matches the UE's power class for the cell's frequency band [50].

It can be observed that the measured UE powers primarily lie within the model-predicted power range. Yet, several individual measurements exceed the predicted maximal power. Apart from model inaccuracy, this may also be caused by retransmissions due to fading effects resulting from unintended UE movements. Because specific Tx power values are not available, the measurements cannot be compared with detailed model predictions.

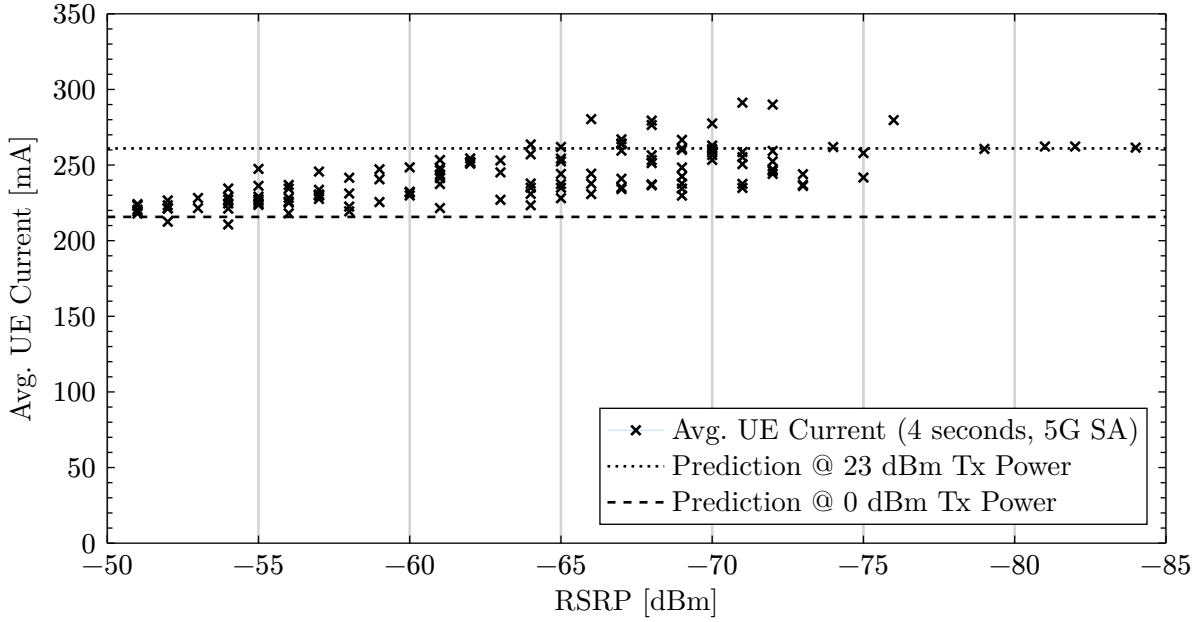


Figure 21: Avg. UE Power by RSRP at 100 Packets per Second

[23]Winter poses unique challenges for outdoor experiments. Figure 21 combines data from two experiments on separate days.

To conclude the validation section, based on the conducted experiments, the developed traffic power model has been proven to conceptually capture the power impact of its parameters, although the specific effect sizes are not accurate and vary depending on the UE and RAN configuration. Further measurements with varying RAN parameters, especially BWP sizes, would be required to evaluate the model in greater detail. Yet, this cannot easily be achieved with commercial 5G deployments. Furthermore, the validation performed in this section is limited by the lack of access to the UE's BWP size and Tx power values. While more detailed measurements – including parameter variations such as higher data rates – can still be achieved with the current setup, an in-depth analysis is not within the scope of this thesis. If more accurate predictions for a given UE and RAN are required, the developed power model may also be adopted via scaling of individual power values to better represent a specific UE. Nevertheless, without further changes, the obtained model predicts the effects of relevant system-level parameters on UE power consumption with sufficient accuracy to be used for UE power estimation in system-level simulations.

4 Handover Optimization

As described in Chapter 1, handover decisions affect QoE via data rates and UE power consumption. Therefore, optimizing handovers can improve QoE. The related work presented in Chapter 2 (Section 2.1) has identified suboptimal cell selection as a frequent cause for reduced UE data rates. In this regard, RL has been used to optimize cell selection for data rates, yet not considering QoE. An exception to this is the QoE-maximizing RL-based cooperative multi-point optimization by Schneider *et al.* [8]. Nevertheless, this does not explicitly determine a UE's serving cell and therefore does not consider handovers. Furthermore, none of the available solutions evaluate the impact on UE power consumption. Therefore, in this section, an RL-based handover optimization approach is developed that specifically considers QoE while taking UE power consumption into account.

Due to the advantages of simulation-based RL prototyping (as motivated in Chapter 1, Section 1.2) and the inaccessibility of a practical 5G environment for experimentation, the work in this chapter relies on a system-level cellular network simulator. In the following section, the development of this simulator will be detailed. Subsequently, Section 4.2 investigates the application of RL to the developed simulation environment and evaluates the impact of the resulting policies on QoE, data rates, and power consumption.

4.1 Simulation Environment

The work of Schneider *et al.* [8] is based on a simplistic system-level cellular network simulator developed in Python. The authors have termed their simulation environment `mobile-env` and published it as open-source software under the permissive terms of the MIT license [43]. Thus, unlike the closed-source simulation environments of several RL-based handover optimization approaches [20], [21], it can be freely applied and modified in related research. Schneider *et al.* study a mobile radio network environment that is strongly related to the cellular handover environment considered in this thesis. Therefore, `mobile-env` may serve as a baseline for the simulation environment. In line with this approach, the following section describes the design and simulation model of `mobile-env`. Subsequently, the adaption of `mobile-env` for this thesis is addressed in Section 4.1.2.

4.1.1 `mobile-env`

`mobile-env` is an abstract, system-level mobile radio network simulator. It supports configurable simulation scenarios, where each scenario represents a fixed number of BSs and UEs that are positioned on a two-dimensional map. An example scenario with two BSs and five UEs is depicted in Figure 22. The positions of the BSs are fixed, while the UEs

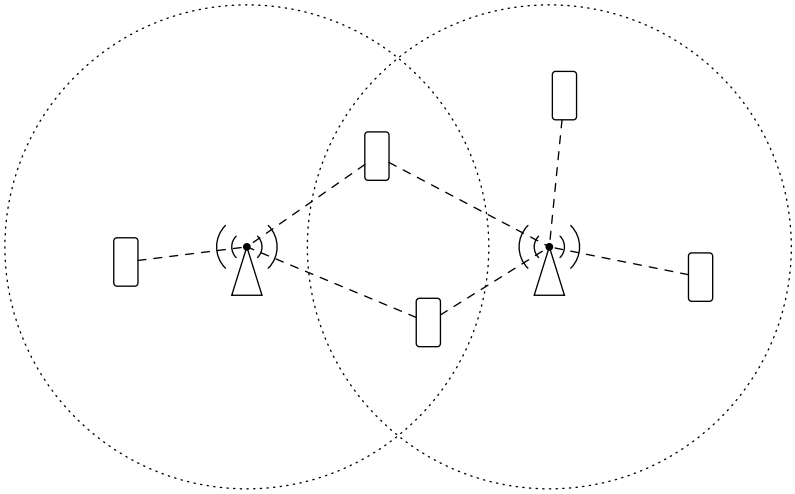


Figure 22: Example `mobile-env` Scenario (Two BSs, Five UEs)

move across the map according to a configurable movement pattern. By default, a random-waypoint movement is used where each UE repeatedly chooses a random point on the map and moves to it with a constant (configurable) velocity. Each UE can connect or disconnect to the various BSs in its reach and – in line with coordinated multi-point communication – can be served by multiple BSs simultaneously. In Figure 22, connections are represented by dashed lines and the maximum range of each BS is indicated by a dotted circle.

Notably, the simulation makes the simplifying assumption that cells are omnidirectional and each BS only operates one cell. Furthermore, no distinction between UL and DL transmissions is made and UEs are assumed to have unlimited data demands[24]. The data rate between a UE and a BS is determined as follows. First, a channel model (Okumura-Hata by default) is used to calculate the distance-based path loss between the UE and the BS. Subsequently, the SNR of the BS's signal at the UE is computed by subtracting the calculated path loss from the BS's fixed transmit power and relating it to the UE's thermal noise level. Inter-cell interference is not considered. Based on the SNR, the Shannon–Hartley theorem is applied to obtain the UE's maximum channel capacity. Finally, the BS schedules its resources to all connected UEs. By default, a resource-fair scheduling is assumed, i.e., the data rate for each of N UEs is calculated as a fraction $\frac{1}{N}$ of the UE's maximum channel capacity. When a UE is connected to multiple BSs, the UE's total data rate is derived as the sum of all BS-scheduled data rates.

`mobile-env` has been designed specifically for RL research. For this reason, it has been implemented as a Gymnasium[25] environment. Gymnasium is a Python library which

[24]The given assumptions may represent a file download scenario when core network speed is no limiting factor.

[25]https://github.com/Farama-Foundation/Gymnasium

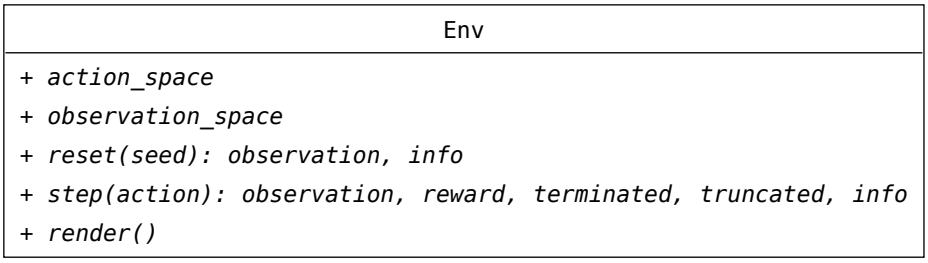


Figure 23: Gymnasium Environment Interface [7]

defines an interface for agent-environment interactions in RL setups [7]. It achieves this by providing an environment base class that can be extended by specific RL environment implementations. The interface of Gymnasium's environment class is outlined as a Unified Modeling Language (UML) diagram in Figure 23. As depicted in Figure 10, an RL agent interacts with its environment by repeatedly supplying an action to it and, in return, obtaining an observation and a numeric reward. To communicate the shapes of expected actions and returned observations to the RL agent, a Gymnasium environment exposes the definitions of action and observation spaces via its `action_space` and `observation_space` properties. By using these definitions, RL algorithms of various compatible frameworks can be trained with Gymnasium environments without any configuration overhead.

RL algorithms base their actions on observations and rewards. To initialize the environment and obtain an initial observation, Gymnasium environments implement a `reset()` method. This method can optionally be supplied with a random number generator seed to make the environment's behavior deterministic and therefore reproducible. In subsequent interactions, the environment's `step()` method is used repeatedly to supply an action to the environment and retrieve the corresponding observation and reward. Additionally, the `step()` method also returns the boolean flags `terminated` and `truncated`, indicating whether the task has reached a natural end or has been truncated (as eventually required for continuing tasks), respectively. Finally, Gymnasium environments can be visualized in each step if they implement the `render()` method. Rendering enables researchers to visually observe an agent's interactions with the environment and to intuitively assess learned policies.

As a subclass of Gymnasium's `Env` class, `mobile-env` implements the core simulation logic in its `step()` method. This involves applying an action (i.e., connecting/disconnecting UEs to/from BSs), moving the UEs according to their configured velocities, computing the updated channel capacities for each UE-BS connection, scheduling the resources of each BS between the connected UEs, deriving each UE's aggregated data rate, and finally assembling an observation vector and computing a reward value. In addition to

Gymnasium's required environment methods, `mobile-env` also implements the `render()` method in a similar manner as depicted in Figure 22.

4.1.2 Environment Adaption

This section elaborates on the steps that have been involved in adapting `mobile-env` to a UE-power-aware cellular handover simulation environment for the application of RL to handover optimization in this thesis. To be applicable in this work, the environment has to (1) restrict each UE to a single serving cell and (2) integrate the developed system-level power model. While the general assumptions of the `mobile-env` simulation do not enable the same level of detail as the simulations by Yajnanarayana *et al.* [20] and Mollel *et al.* [21], implementing a more realistic physical simulation model is out of the scope of this thesis. Based on the object-oriented architecture of `mobile-env`, the adaption was therefore planned as two actionable items:

(1) Create a subclass of the `mobile-env` environment that overrides the simulation's `apply_action` method to limit each UE to one BS at a time. Specifically, this can be achieved by disconnecting the UE from the active BS when the supplied action demands a handover to another BS.

(2) Create a `UserEquipment` subclass that adds a method to retrieve a UE's current power consumption based on the power model from Chapter 3. Regarding the model's input parameters, the DL BWP size B_D is known from the configured BS bandwidth in the `mobile-env` simulation. Yet, all other model parameters are not explicitly simulated by `mobile-env`. Therefore, assumptions have to be made for the remaining model parameters, as detailed below.

- **DRX**: The operator-defined DRX parameters depend on a UE's 5QI. In a general web traffic 5QI, DRX parameters similar to those observed in Section 3.3 can be assumed ($T_C = 320$ ms, $T_I = 100$ ms, $T_{\text{On}} = 20$ ms).
- **Tx Power**: The UE's Tx power can be estimated based on the simulation environment's SNR values. Recent measurements by Jörke *et al.* [51] have shown a linear trend between a UE's SNR and its RAN-configured Tx power. The trend can be formulated as $x(\text{SNR [dB]}) = 23 \text{ dBm} \cdot \left(1 - \frac{\text{SNR}}{40}\right)$ for $\text{SNR} \in [0, 40]$.
- **Mean UL/DL Inter-Arrival Times (IATs)**: The average packet IAT depends on a UE's application and is not simulated in `mobile-env`. However, based on the assumption of data-demanding UEs, DL file transfers using TCP may be a plausible application. The IATs of TCP packets depend on the Round-Trip Time (RTT) between the sender and receiver of a data stream. Assuming that

the RTT does not vary significantly, using a fixed average IAT for the power model is plausible in the considered environment. Varying payload data rates can be expected to affect TCP's packet sizes without significantly impacting IATs. Motivated by the power model's accurate validation results at 100 packets per second, a value of $T_{P_D} = T_{P_U} = 10$ ms will be assumed.

While the implementation of Item (1) resulted in the desired changes, during the implementation of Item (2), highly unrealistic SNR values in the order of −76 dB were encountered. The process of locating and solving the underlying issues in the physical simulation of `mobile-env` was accompanied by several iterations of refactoring in which the data model was adapted to enable the implementation of Item (2). Specifically, the simulation state in `mobile-env` is primarily stored in environment-global dictionary objects and handled by environment methods. This requires extensions to be implemented globally at the environment level and prohibits fine-grained object-oriented model extensions as planned in Item (2). While `mobile-env` follows an object-oriented approach, model objects like `BaseStation` and `UserEquipment` primarily hold static parameters (e.g., BS bandwidth, UE velocity) while their simulated state (e.g., connections, scheduled data rates, SNR) is maintained globally by the environment.

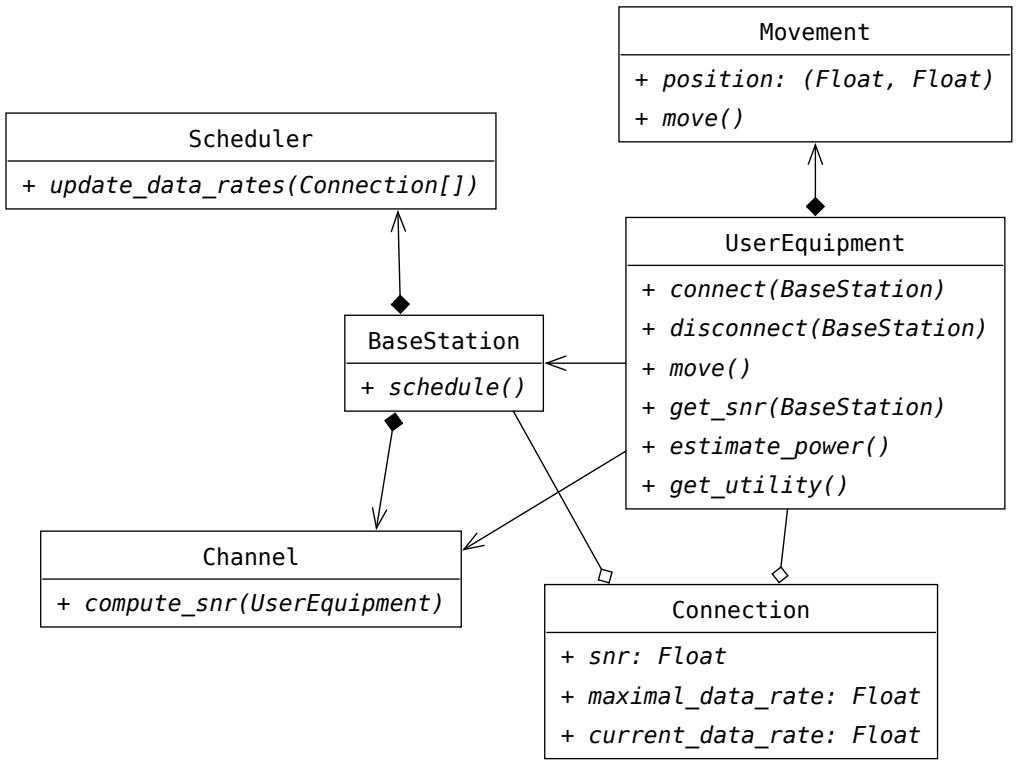


Figure 24: Adapted Simulation Data Model

The refactoring process involved in identifying simulation inconsistencies ultimately resulted in a more object-oriented, decentralized data model, as detailed in Figure 24. By introducing the new data model, the complexity of the environment's `step()` method has been reduced and the readability and extensibility of the simulation environment have been improved. In the `step()` method, the environment now invokes each UE's `move()` method, causing the UE to change its position (maintained by its `Movement` object) and recompute the SNR and `maximal_data_rate` (channel capacity) of each of its `Connection`s. Subsequently, the environment invokes each BS's `schedule()` method, thereby updating the `current_data_rate` of each connection. The adaption of the presented data model has reduced redundant SNR computations via caching and the overall simulation speed has been improved by a factor of seven in a two-BS, five-UE scenario. Ultimately, the new data model has enabled the integration of the UE power model as detailed in Item (2). The resulting environment is available online in the repository of this thesis and may be used for future upstream contributions.

4.2 Reinforcement Learning

Based on the UE-power-aware cellular handover simulation environment developed above, this section examines the application of RL algorithms to optimize QoE via data rates and UE power consumption. While off-the-shelf RSRP-driven cell selection does not necessarily yield optimal data rates [13], [14], [20], the power model developed in this thesis suggests that – within the domain of 5G NR cells – it may nevertheless yield an optimal power consumption due to low transmit powers. This is underlined by the power measurements in Figure 21, which indicate reduced power consumption for higher RSRPs. Therefore, by deviating from RSRP-based handover decisions, an RL-driven handover algorithm may unintentionally increase UE power consumption, despite optimizing QoE as a function of data rates and UE power. An example for a situation with conflicting data rate and power consumption is provided in Figure 25, where the UEs are distributed unevenly across the two cells. When the leftmost UE stays connected with the left cell, all UEs can achieve higher data rates than by sharing a single cell, yet the leftmost UE has to use increased transmit power. By comparison, an RSRP-based handover mechanism would already have handed the leftmost UE over to the right cell.

Applying the developed UE-power-aware cellular handover simulation environment allows to jointly evaluate and compare the effects of an RL agent's decisions on data rates and UE power consumption. This section is organized in three parts, as follows. Section 4.2.1 handles the definition of actions, observations, and rewards. Subsequently,

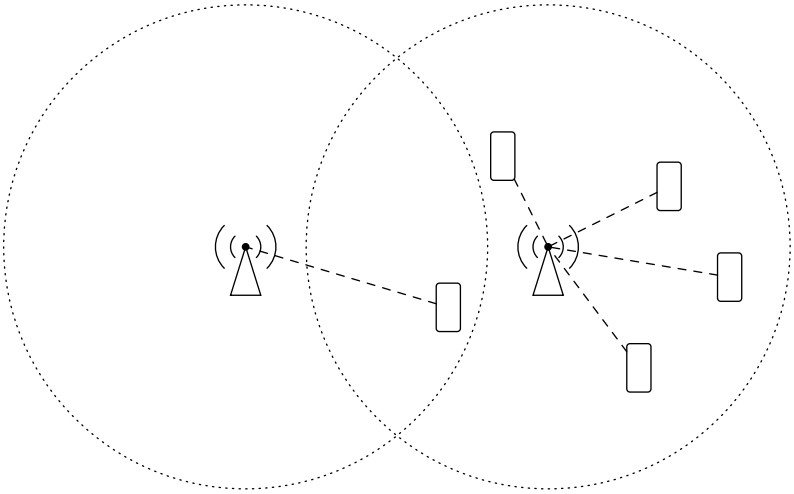


Figure 25: System State Example for Conflicting Data Rate and Power Consumption

incremental improvements to optimize the RL agent's performance are discussed and applied in Section 4.2.2. Finally, the best policy is compared against a conventional RSRP-driven handover approach in Section 4.2.3.

4.2.1 Interface

In Chapter 2, three distinct RL-based cell selection solutions have been presented – among them two data-rate-driven handover optimizations [20], [21] and one QoE-driven cooperative multi-point solution [8]. Despite the different scope of each approach, a common baseline is their observation space. Notably, all three approaches rely on an observation vector that, for each UE, contains the measured signal strength (either RSRP or SNR) of each cell. This data is readily available at the RAN, since the UEs provide it via their RRM measurement reports. The combination of all cell SNRs for a given UE allows an RL agent to estimate the UE's physical location and therefore take appropriate handover actions. Furthermore, Mollel *et al.* [21] and Schneider *et al.* [8] also include one-hot-encoded connection vectors for each UE in their observation space, informing the agent about the current cell choices. Based on the promising results of the presented approaches, the same observation space will initially be applied in the next section's RL experiments.

To consider QoE as a function of data rates, Schneider *et al.* rely on a logarithmic UE utility function, inspired by empirical research indicating such effects [22], [23]. The function is included in `mobile-env` and depicted in Figure 26. Motivated by the observations of Fiedler *et al.* [22], it will likewise be used as an estimation for QoE based on the simulated data rates in the following RL experiments. For consistency with the approach of Schneider *et al.*, each UE's QoE value will also be included in the observation. Similarly, the mean UE QoE will serve as the reward signal. Finally, the agent's action space is

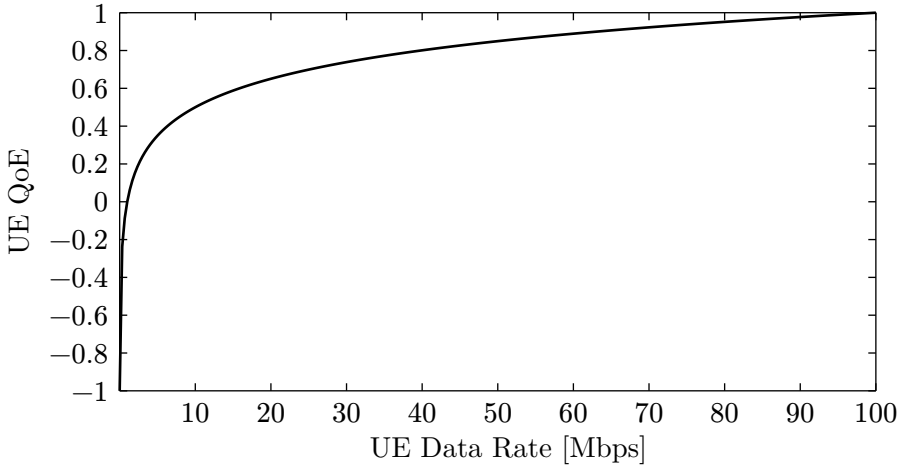


Figure 26: Logarithmic UE Utility Function from [8]

chosen as a vector of cell ids, where each vector element represents the selected cell of a specific UE.

4.2.2 Evolution

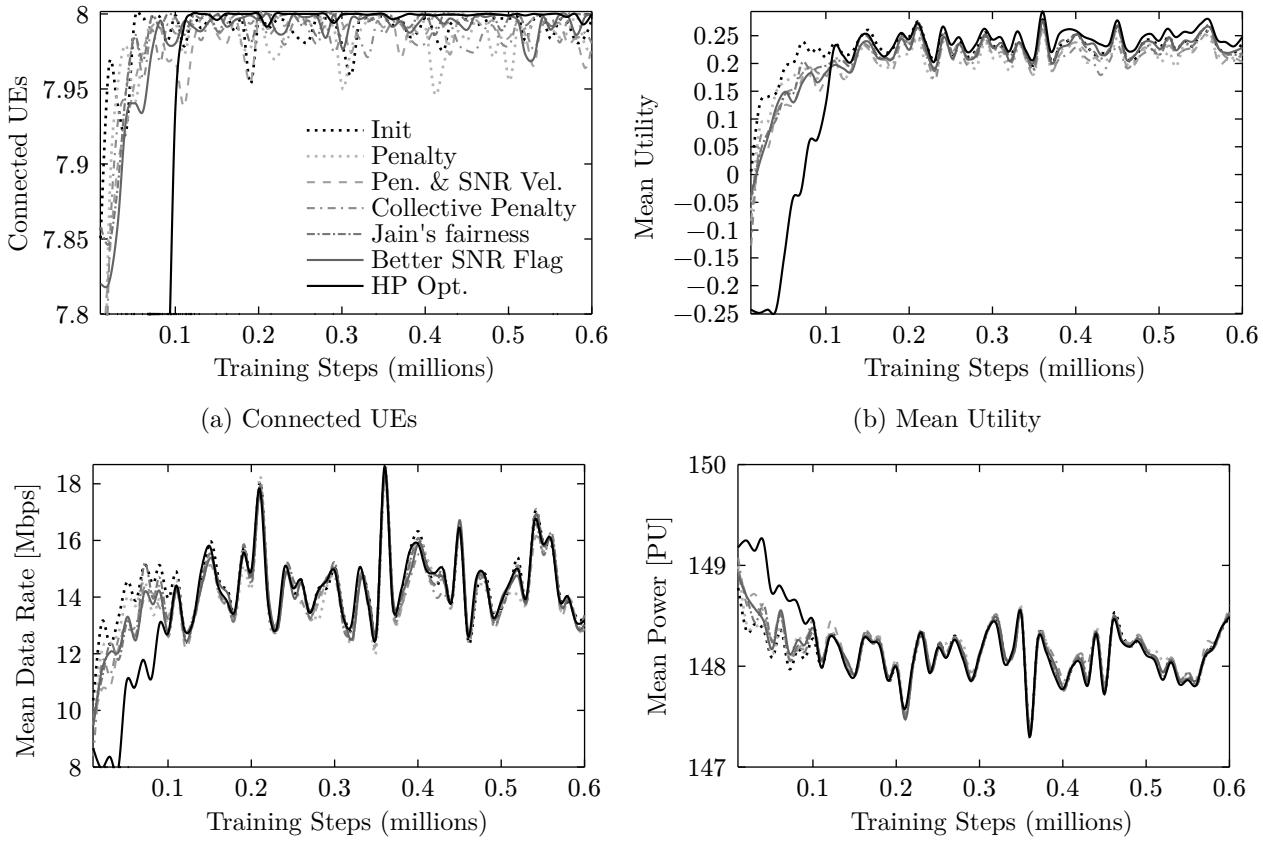
To experimentally investigate different reward functions and observation space extensions, a series of training sessions with varying configurations has been carried out. In each session, a policy is learned using the state-of-the-art RL algorithm PPO [36][26] in a small simulation environment comprising two BSs and eight UEs. Similar to the experiments by Schneider *et al.*, each episode is truncated after 100 time steps. All UEs are configured to move with a velocity of 15 meters per time step. The environment simulates an area of 2.6×1.6 km, which is fully covered by the two cells, yet contains cell-edge regions for each cell. The horizontal distance between the BSs is 1.2 km and the BSs are centered on the map. Both cells are operated on the 5G NR band n78 with a bandwidth of 90 MHz and a center frequency of 3600 MHz. The BS und UE height are set to 15 m and 1.5 m, respectively. Resource-fair scheduling is applied and the BSs' transmit power is configured to 30 dBm. A standard thermal noise density of -173.8 dBm/Hz is assumed for UE receptions. During the training sessions, every 100 trained episodes, the agent's policy has been evaluated for a duration of 20 episodes. Evaluation metrics include the average number of connected UEs, the mean data rates, QoEs, and connected UE power consumption. The following paragraphs detail the challenges and solutions involved in optimizing QoE via handover decisions.

Notably, training a PPO agent for 0.6 million training steps with observation and rewards configured as described in the previous section already yields an intelligent policy

[26]The PPO implementation of Stable Baselines 3 was used in this work.

that significantly improves the evaluation metrics over the course of the training period. This can be observed via the dotted "Init" curves in Figure 27. For each evaluation metric, the average value in an evaluation period has been plotted as a curve over the training step count. A single evaluation period is only 20 episodes (i.e., 2000 steps) long. Therefore, the average values in each evaluation period are subject to noise caused by varying UE movement patterns in each evaluation period. Namely, in one evaluation, UE movements may lead to shorter average BS-to-UE distances and therefore enable higher data rates than in another evaluation. For easier comparison of the curves in Figure 27, the lines have been interpolated.

Despite the agent's improvement on the considered metrics, the obtained policy is not yet ideal. Specifically, when observing the policy using the environment's rendering feature, a number of flaws can be identified. In the following paragraphs, each of these flaws is presented and solutions are developed by adapting the reward function and observation space. The training evaluation curves for each adaption are included in Figure 27, respectively.



(a) Connected UEs

(b) Mean Utility

(c) Mean Data Rate [Mbps]

(d) Mean Power [PU]

Figure 27: Training-Period Evaluation Metric Averages over Training Steps

Repeated Handovers: When a UE is within the range of two cells, the obtained initial policy occasionally shows an alternating handover behavior where a UE is repeatedly handed over between two cells. While this may increase the agent's total reward by equally sharing cell bandwidths, it must be prevented in practice to avoid service interruptions for the UE. A natural approach involves penalizing handovers, as applied by Mollel *et al.* [21]. Therefore, for each handover action in a step, the QoE in that step is set to -1. In successive steps without handovers, the QoE values remain unchanged. Intuitively, this represents the drop of a user's QoE during a handover where application traffic is interrupted for a short time.

By applying this change (denoted as "Penalty" in Figure 27) to the reward function, the amount of repeated handovers is reduced. Yet, repeated handovers still occasionally occur shortly before reaching a cell edge. On closer examination, it becomes evident that the agent can only infer a UE's position from its SNR measurements, yet cannot know the movement direction and velocity of the UE. Therefore, the agent cannot be certain whether a UE at the cell edge has just arrived in the cell or is about to leave the cell. A solution to this is the inclusion of an SNR velocity vector in the observation space. Due to the variance of SNR values in real systems, this velocity vector would likely require the use of a moving average to avoid strong influence of temporary SNR fluctuations. Yet, the movement and channel model in the `mobile-env`-based simulation environment do not exhibit short-term variations. Therefore, in the context of this work, an SNR velocity vector can be computed as the difference of the previous and the next SNR vector. Adding a per-UE velocity vector to the observation space finally solves the repeated handover issue. Notably, SNR velocities alone, without penalizing handovers, are not sufficient to prevent repeated handovers.

Starvation: It can be observed that, occasionally, in loaded cells, the agent does not connect an individual UE. This may optimize the average QoE under certain conditions, by "sacrificing" an individual UE to avoid reducing the QoE of all other UEs in the cell. Yet, this behavior is inacceptable in practice. The issue can be solved by penalizing the agent with the maximum negative reward (-1) in any step in which at least one UE is not connected, but in reach of a BS. The resulting evaluation curves are listed as "Collective Penalty" in Figure 27.

Fairness: Another issue with the obtained policies is that, in some situations, they are inherently unfair, e.g., by excluding a UE from a loaded cell and connecting it to a neighboring cell to optimize overall UE QoE, as depicted in Figure 25. Apparently, optimizing mean QoE and optimizing QoE fairness pose a trade-off in certain situations.

To control the amount of desired fairness – at the cost of overall QoE, the mean QoE reward signal can be weighted with a dedicated fairness reward signal. A commonly used fairness metric is "Jain's fairness index" [52]. To improve QoE fairness, Jain's fairness index can be computed, normalized to the reward range $[-1, 1]$, and subsequently be weighted with the mean QoE reward signal. Several training sessions with various fairness weights have shown a fairness weight of 0.5 to yield the best results with respect to the average UE QoE in evaluation episodes.

Missed Handover Chances: Based on the given observation space, the agent sometimes waits to apply handover actions that, intuitively, seem to optimize both mean QoE and QoE fairness. In case the agent misses information to consider a specific UE's immediate handover, it can be helpful to include a flag in the observation space that indicates when a better choice may be available for the given UE. To evaluate this approach, each UE's observation vector has been extended with a flag that is set to `true` as soon as a cell with a higher SNR than the serving cell is available. Notably, this mechanism – listed as "Better SNR Flag" in Figure 27 – has further improved the agent's reward.

Finally, to further optimize the results of the RL agent, the hyper parameter tuning framework Obtuna[27] has been applied to automatically optimize the hyper parameters of the PPO algorithm for increased rewards. Primarily, this has shown that a decreased learning rate (approximately 0.00008, as opposed to Stable Baselines' default value of 0.0003) yields better results for the given task. In Figure 27, the final and best result obtained by combining all above methods with tuned hyper parameters is denoted "HP Opt." and plotted as a black line.

4.2.3 Evaluation

With regard to the training evaluation metrics from the previous section, a broad evaluation and comparison among the various RL policies has already been obtained. Contrary to the previous approach, this section only considers the last and best policy and evaluates it in greater detail. While the evaluation metric averages in Figure 27 were obtained by applying each policy for only 2000 steps, the evaluation in this section collects the metrics over a duration of one million simulation steps to reduce the impact of random UE movements.

Table 9 details the average metrics of the evaluated RL-based handover policy in direct comparison with the values obtained by an off-the-shelf RSRP-based handover algorithm. Both approaches have been tested with an identically seeded environment. Therefore, the

[27] https://optuna.org/

Table 9: Average Evaluation Metrics for RL vs. RSRP-based Handovers

Metric	RSRP Mean (std.)	RL Mean (std.)	Difference [%]
UE QoE $\in [-1, 1]$	0.242 (0.138)	0.244 (0.135)	+0.786 %
UE QoE Fairness Index (Jain)	0.919 (0.049)	0.921 (0.046)	+0.196 %
UE Data Rate [Mbps]	14.707 (12.673)	14.486 (12.077)	−1.507 %
# Connected UEs	8.000 (0.000)	7.999 (0.028)	−0.01 %
Connected UE Power [PU]	148.078 (2.314)	148.082 (2.312)	+0.002 %

comparison involves no movement-induced noise. It can be observed that both algorithms do not yield significantly different results in the given simulation environment. Yet, the RL policy has managed to achieve a slightly better UE QoE and fairness of QoE, at the cost of a minor average data rate reduction. The improvement of QoE, however, only totals less than one percent of the RSRP-based QoE average. With respect to UE power consumption, no significant relative difference can be observed.

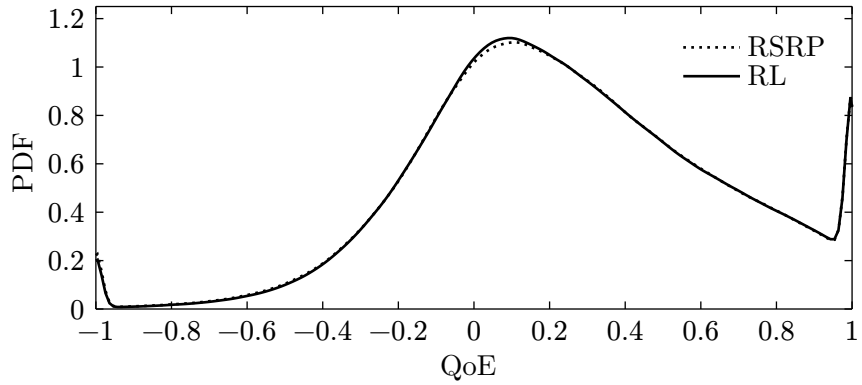


Figure 28: Comparison of QoE Probability Density Functions

To enable an in-depth analysis of the achieved QoE improvement, Figure 28 plots the probability density function of all UEs' QoE values throughout the simulation. While the overall distribution does not vary significantly between the two approaches, the RL-based approach has slightly reduced the probability of low QoE values and instead concentrated it in the medium QoE region. In the following chapter, the obtained results will be discussed in the broader context of this thesis.

5 Discussion

The work in both topics of this thesis – UE power prediction and RL-based handover optimization – has led to a number of findings which are summarized in Section 5.1. Based on these findings, options for future work are discussed in Section 5.2.

5.1 Key Findings

The development of a system-level UE power model in Chapter 3 has enabled a detailed understanding of the factors influencing UE power in 5G NR deployments. Most importantly, UE power is significantly impacted by the network-configured bandwidth part. Intelligent dynamic adaption of a UE's BWP can result in major UE power savings. Furthermore, a large amount of UE power is consumed solely for staying connected, more specifically, to monitor the radio channel for transmissions by the BS. Only a comparably small fraction of UE power can be saved by optimizing cellular handovers. Notably, these power savings only rely on the varying transmit powers required for different cells. By contrast, the optimization of RAN parameters has a far larger impact on UE power consumption.

With respect to handover optimization for QoE, the diversity of radio environments has a significant impact on the performance gains enabled by RL. Specifically, related work using more realistic physical channel simulations for RL-based handover optimization has consistently outperformed traditional handover algorithms, despite involving similar observation spaces as the approaches in this thesis. In comparison, the RSRP-based handover algorithm used for the evaluation in Section 4.2.3 has, without further ado, achieved close-to-optimal results in the simplistic simulation environment of this thesis, leaving not much room for improvement. The comparably weak performance gain of the obtained RL policy is comprehensible, considering that a key strength of RL approaches lies in the exploitation of subtle environment dynamics that are not captured by many conventional algorithms. Yet, applying RL-based handovers to more diverse radio environments – such as the ones observed by Hassan *et al.* [13] – may be an impactful step in closing the gap between practically observed maximum data rates and those achievable by optimal cell selection.

5.2 Future Work

Several topics qualify for future work. In regard to power modeling, the system-level power model developed in this thesis – despite yielding several promising results in the measurements performed in Section 3.3 – does not accurately predict UE power. Specifically, it

is based on vastly simplified assumptions regarding physical resource scheduling. While a further in-depth theoretical refinement of the developed model may not significantly improve power predictions, the observed effects of the model's input parameters on UE power consumption may form the foundation for a hybrid approach. Namely, a simplified parameterized mathematical model may be derived by generalizing the theoretical model and – in a similar approach to the empirical power modeling of Jensen *et al.* [6] – may be fitted with measured power values. This way, the various effect sizes of model parameters on the power prediction may be captured more accurately.

Concerning the cell-selection-induced data rate gap in today's mobile radio networks, network operators may profit from an online learning solution that can be deployed in existing mobile networks. Most RL networking research relies on network simulations. Yet, as observed in Chapter 4, despite carefully modeling relevant system dynamics, simulations may deviate significantly from the behavior of real networks. By utilizing the O-RAN platform, a hybrid online-learning RL-based handover solution may be developed that learns policies based on the data retrieved from a real network. To avoid service disruptions by such an application, a conventional control algorithm might serve as a safety measure to control disruptive actions. Ultimately, such a system may yield more impactful results than abstract simulation-based RL solutions, which need to be carefully evaluated with respect to their applicability to real systems.

6 Conclusion

This thesis has investigated the optimization of handover decisions with respect to the Quality of Experience (QoE) perceived by users of mobile radio network devices. QoE depends on many factors, including data rates and power consumption of battery-powered mobile devices (UEs). To study the impact of handover decisions on UE power consumption, a theoretical system-level UE power model for 5G New Radio networks has been developed. The model has been derived from an existing fine-grained industry-standard power model and validated with UE power measurements obtained in a commercial 5G deployment in cooperation with the mobile network operator Deutsche Telekom AG. While the model's predictions properly represent the effects of system-level parameters on UE power consumption, the predicted effect sizes differ from the effect sizes that were observed in the power measurements. Nevertheless, the developed power model is extensible and may be adapted to improve its accuracy for specific UEs if more precise power predictions are required in a given network configuration. For practical use in handover optimization, the developed power model has been implemented as a Python package and integrated into an existing system-level cellular network simulator.

To optimize handover decisions for QoE with respect to data rates and UE power consumption, Reinforcement Learning (RL) has been applied to the developed UE-power-aware network simulator. Several approaches have been investigated to improve the performance of RL in handover optimization. The resulting handover decision policy has been evaluated and compared with an off-the-shelf handover mechanism, revealing a minor QoE improvement. Finally, the developed power model, measurement scripts, and RL-tailored UE-power-aware handover simulation environment are openly available[28] under the MIT license. They can be built upon in future research to further enhance QoE in mobile radio networks.

[28] https://github.com/bjoluc/5g-handover-optimization

A Stochastic DRX Sleep Ratio Analysis

This section explores analytical approaches to derive the DRX-induced sleep ratio r_{sleep} based on the packet inter-arrival times T_{P_D} and T_{P_U} in combination with the DRX parameters T_C, T_I, and T_{On}, as listed in Table 8.

A1 Expectation-Based Approach

Let t_a and t_s be random variables, where t_a denotes the duration of a DRX active period and t_s denotes the time the UE spends in a sleep state between two active periods. The ratio of sleep time can thus be expressed as

$$r_{\text{sleep}} := \frac{\mathbb{E}[t_s]}{\mathbb{E}[t_a] + \mathbb{E}[t_s]} \tag{17}$$

To calculate r_{sleep}, subsequently, the expectation values for t_a and t_s are derived based on the input parameters. Assume that at least one packet is transmitted at the beginning of a DRX active period[29]. This can be either an UL packet which arrives in a DRX sleep period and resets the inactivity timer[30], or a (possibly buffered) DL packet arriving at the beginning of a DRX On duration.

Every packet arrival (UL or DL) after the first transmission resets the inactivity timer and therefore contributes to the active time t_a. Therefore, t_a can be decomposed as the sum of a number of packet inter-arrival times t_{P_i}, each shorter than the inactivity timer T_I, plus one final T_I for timer expiry:

$$t_a = \sum_{i=1}^{N} \left(t_{P_i} \right) + T_I \tag{18}$$

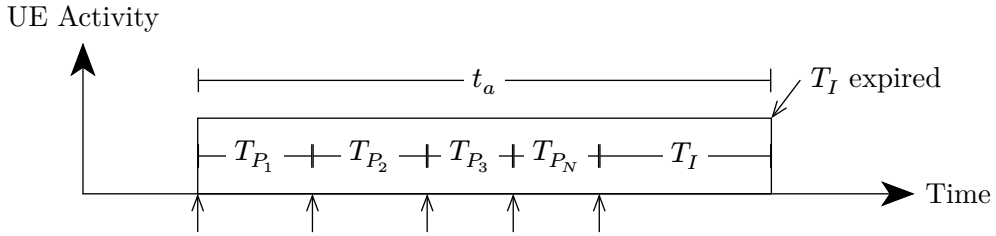


Figure 29: DRX Active Period with UL/DL Packet Arrivals

[29] All packets from the previous sleep duration are received in a batch at the beginning of the On duration and will be considered as one transmission for brevity, assuming the base station schedules as many packets as feasible in one slot.

[30] Technically, it is the UL grant that resets the inactivity timer, but assuming that the scheduling delay does not vary significantly between UL packets, UL grants can be considered to have the same probability distribution as UE UL packet arrivals.

Figure 29 depicts this scenario with $N = 4$ packet arrivals (marked by arrows on the time axis) after the initial transmission at the beginning of the active period. Both UL and DL packet arrivals are modeled as Poisson processes. Let $A_U(t)$ and $A_D(t)$ be random variables denoting the number of UL packet arrivals at the UE and DL packet arrivals at the BS, respectively, during time t:

$$A_U(t) \sim Poisson\left(\frac{t}{T_{P_U}}\right),\ A_D(t) \sim Poisson\left(\frac{t}{T_{P_D}}\right) \tag{19}$$

Due to the properties of the Poisson distribution, the number of UL and DL packet arrivals $A(t)$ in time t is known to be Poisson-distributed as well.

$$A_U(t) + A_D(t) =: A(t) \sim Poisson\left(\frac{t}{\lambda}\right) \text{ with } \lambda := \frac{1}{\frac{1}{T_{P_U}} + \frac{1}{T_{P_D}}} \tag{20}$$

Inter-arrival times of a Poisson distribution are exponentially distributed with the inverse of the distribution's parameter, i.e., $Exp(\lambda)$ in case of UL/DL packet inter-arrival times. However, this does not fully apply to the series of random variables t_{P_i} which are each known to be shorter than the inactivity timer T_I. Therefore, the distribution of t_{P_i} is truncated at T_I:

$$t_{P_i} \sim TExp(\lambda, T_I) \tag{21}$$

The expectation of a right-truncated exponential distribution can be derived by conditioning on the probability that a value lies within the truncation range:

$$\text{Let } X \text{ be a r.v. with } X \sim Exp(\lambda).$$

$$\mathbb{E}[X \mid x \leq c] = \frac{\int_0^c x\lambda e^{-\lambda x}}{\mathbb{P}(X \leq c)} = \frac{\frac{1-e^{-\lambda c}(\lambda c+1)}{\lambda}}{1-e^{-\lambda c}} = \underbrace{\frac{1}{\lambda}}_{\mathbb{E}[X]} \cdot \underbrace{\frac{1-e^{-\lambda c}(\lambda c+1)}{1-e^{-\lambda c}}}_{\text{truncation factor} \leq 1} \tag{22}$$

Next, the distribution of the random variable N, measuring the number of timer-resetting packet arrivals within t_a, needs to be determined. Due to the memoryless property of the Poisson distribution, the probability for another packet to arrive within T_I after the previous packet's arrival is equal for each packet. Using the random variable $A(t)$ from above, it can be written as $\mathbb{P}(A(T_I) > 0)$. Thus, whether the inactivity timer expires after receiving packet i can be modeled as a Bernoulli trial with the success probability p where

$$p = \mathbb{P}(A(T_I) = 0) = e^{-\frac{T_I}{\lambda}} \tag{23}$$

N is therefore the number of Bernoulli trials needed to gain a "success" outcome (i.e., the inactivity timer expires) and is known to be geometrically distributed:

$$N \sim G\left(e^{-\frac{T_I}{\lambda}}\right) \Rightarrow \mathbb{E}[N] = \frac{1}{e^{-\frac{T_I}{\lambda}}} = e^{\frac{T_I}{\lambda}} \tag{24}$$

Based on Equation 18, we can derive the expectation of t_a as

$$\mathbb{E}[t_a] = \mathbb{E}\left[\sum_{i=1}^{N}\left(t_{P_i}\right) + T_I\right] = \mathbb{E}[N] \cdot \mathbb{E}\left[t_{P_i}\right] + T_I = e^{\frac{T_I}{\lambda}} \cdot \frac{1 - e^{-\lambda c}(\lambda c + 1)}{\lambda(1 - e^{-\lambda c})} + T_I \tag{25}$$

Each DRX active pediod t_a is followed by a sleep period t_s. The duration of the sleep period depends on when the next active period starts. This may either be the case due to an UL packet arrival, or due to the beginning of the next On duration, as demonstrated in Figure 30.

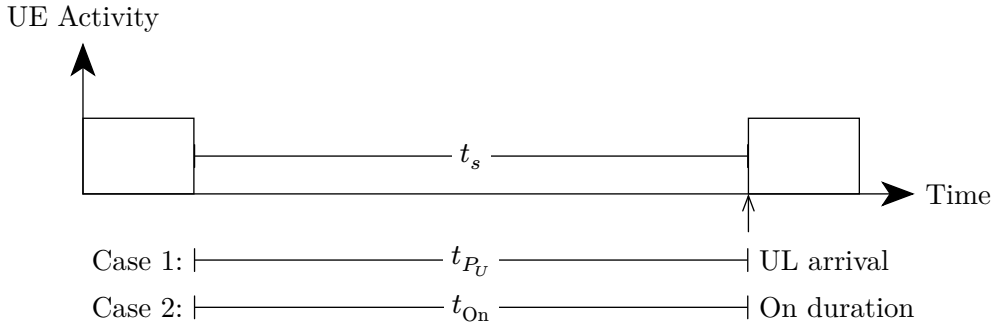


Figure 30: Duration of a DRX Sleep Period

Let the random variables t_{P_U} and t_{On} denote the time from the beginning of the sleep period to the next UL packet arrival or On duration, respectively. The first of these two events ends the sleep period, yielding

$$t_s = \min\left(t_{P_U}, t_{\mathrm{On}}\right) \tag{26}$$

To compute the expectation of t_s, the distributions of t_{P_U} and t_{On} both need to be known. Due to the Poisson distribution of UL traffic, the distribution of t_{P_U} is known to be exponential, depending on the mean UL inter-arrival time T_{P_U}:

$$t_{P_U} \sim Exp\left(T_{P_U}\right) \tag{27}$$

However, the distribution of t_{On} directly depends on the distribution of the moment a sleep period begins within the DRX cycle. Due to the buffering of DL traffic at the BS,

the moment in which a sleep period begins is not uniformly distributed throughout the DRX cycle. Specifically, it is impacted by the duration of the previous sleep period which affects the probability for buffered DL traffic to be transmitted at the beginning of the On duration, therefore introducing recursion. To break this recursion, the beginning of a sleep period might be assumed to be uniformly distributed throughout the DRX cycle. Yet, depending on the input parameters, this may introduce a significant error. While there may be methods to handle the recursive dependency, the complexity of a potential solution is likely out of the scope of this thesis. Therefore, the expectation-based approach of this section is not further explored.

A2 Probability-Based Approach

An alternative approach derives r_{sleep} via the expected active time *per DRX cycle*. This, in turn, can be determined by deriving the probability for the UE to be active at a given point of time in the DRX cycle and integrating that probability over each point of time in the cycle[31]. Specifically, consider a DRX cycle of length T_C. Let $A = \bigcup_i A_i$ be the conjunction of random intervals $A_i \subseteq [0, T_C]$ during which the UE is in DRX active state, i.e., A is a set of time points. Define the length of A as $|A|$. Futher, let $1_A(x)$ be the indicator function of A:

$$1_A(x) := \begin{cases} 0 \text{ if } x \notin A \\ 1 \text{ if } x \in A \end{cases} \tag{28}$$

Then it holds that

$$|A| = \int_0^{T_C} 1_{A(x)} dx \tag{29}$$

Using the properties of expectation and Fubini's theorem [53] (given that $|A|$ is non-negative), the expected active time within a DRX cycle can therefore be written as

$$\mathbb{E}[|A|] = \int_0^{T_C} \mathbb{E}[1_A(x)] dx = \int_0^{T_C} \mathbb{P}(x \in A) dx \tag{30}$$

$\mathbb{P}(x \in A)$ denotes the probability that the UE is active at a given time $x \in [0, T_C]$. During the On duration, i.e., $x \in [0, T_{\text{On}}]$, the UE is always active ($\mathbb{P}(x \in A) = 1$). At any other time $x \in (T_{\text{On}}, T_C]$, the probability $\mathbb{P}(x \in A)$ depends on the events prior to x. More

[31] This idea was brought up by Prof. Dr. Richthammer at the Stochastics Working Group of the Math Department at Paderborn University – thank you, Mr. Richthammer!

precisely, the UE is active at time x if and only if it has woken up at some time in the past and has been kept awake by traffic since then. This observation reduces the derivation of $\mathbb{P}(x \in A)$ to a subproblem which is examined in more detail below.

Consider a time interval $T = [0, t]$, where the inactivity timer has been reset at time zero. We are looking for the probability that the UE remains active during T, i.e., the inactivity timer never expires within T. This is the case if and only if there is no duration of length T_I within T in which no packet arrives. To derive this probability, we will condition the number of arrivals in T to a fixed $N \in \mathbb{N}$ and apply the law of total probability, i.e., sum up the conditional probabilities over all N, weighted with the marginal probability of each N.

Conditioning on a number N of arrivals within T, define the arrival times as $X_1, X_2, ..., X_N$ and the durations between arrivals as $t_1, t_2, ..., t_{N+1}$, as illustrated in Figure 31. Note that $t_1 = X_1$ represents the time from the left limit of T to the first arrival X_1 and $t_{N+1} = t - X_N$ measures the time from the last arrival X_N to the right limit of T.

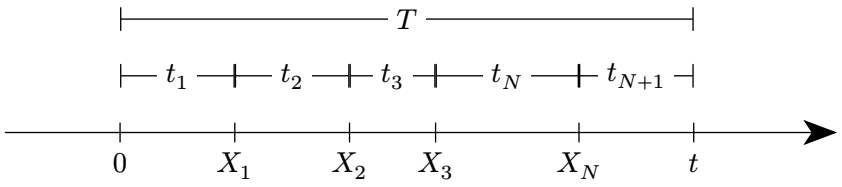


Figure 31: Partition of a Time Interval T

Because the number of arrivals during T is conditioned to N and arrivals are modeled as a Poisson process, the arrival times X_i are independent of each other and uniformly distributed over T. Per definition, the values of X_i are ordered increasingly. Therefore, X_i is the ith order statistic of a set of N independent and identically distributed (i.i.d.) random variables. Order statistics of uniform i.i.d. random variables are Beta-distributed with known parameters:

$$\frac{X_i}{t} \sim Beta(i, N - i + 1) \tag{31}$$

For the first inter-arrival time t_1, this translates to $\frac{t_1}{t} = \frac{X_1}{t} \sim Beta(1, N)$. Intuitively, all scaled t_i are expected to follow the same distribution as t_1, although they are constrained to sum up to one and thus cannot be independent. It is a proven fact [54] that the scaled IATs in this situation follow a flat $Dirichlet(\{1, ..., 1\}, N + 1)$ distribution, which is a multivariate generalization of the Beta distribution. This confirms that each t_i is $Beta(1, N)$-distributed. Given values for each scaled inter-arrival time (where all IATs sum

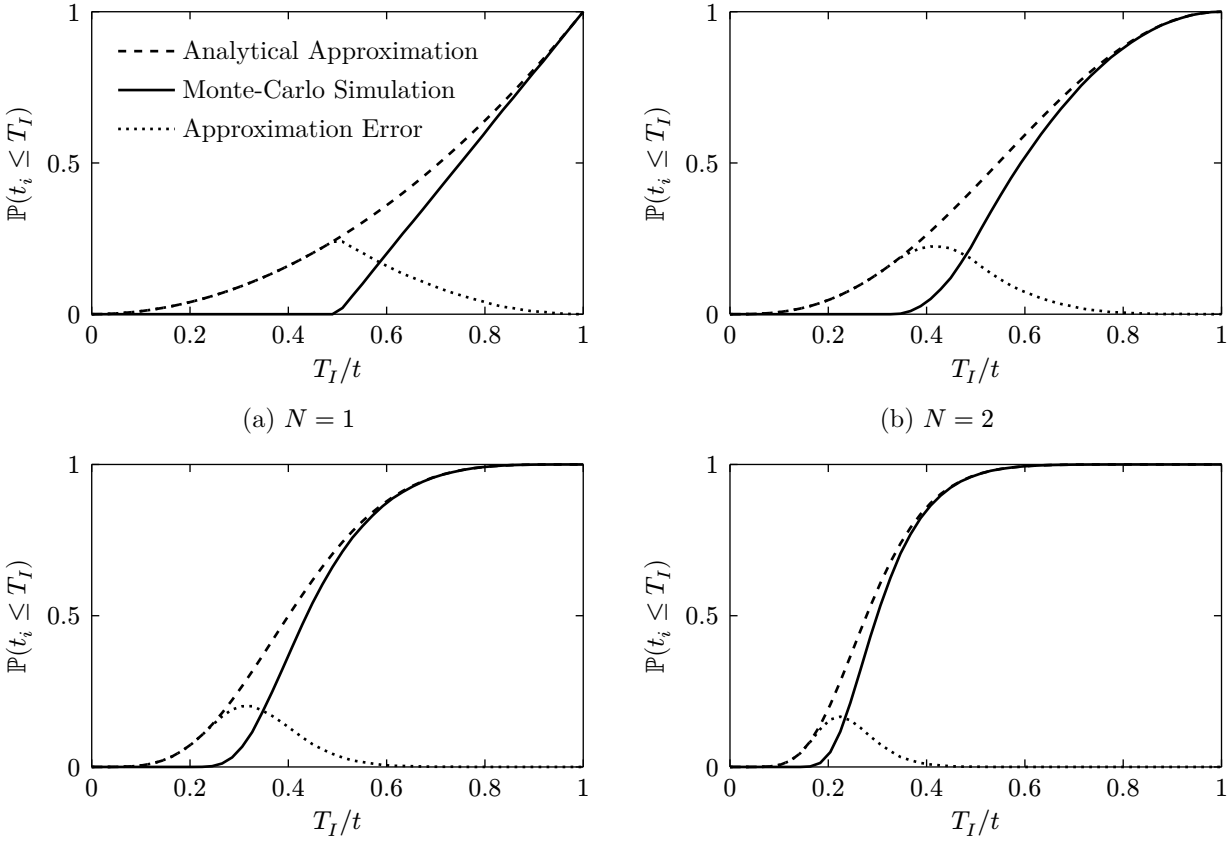
up to one), the probability density function of the flat Dirichlet distribution serves as the joint density function for all IATs.

The probability of interest can be expressed as $\mathbb{P}(\max\{t_1, ..., t_{N+1}\} \leq T_I)$. Deriving the exact probability requires integrating over the Dirichlet distribution's density function for all combinations of t_i that satisfy the sum constraint and have at least one $t_i > T_I$. An analytical solution to this is complex and out of the scope of this thesis, considering that the problem at hand is only a subproblem, the solution of which itself needs to be summed up over N and integrated in a later step. Instead, in the following paragraphs, a simple approximation will be explored and evaluated by comparing it to a Monte Carlo simulation. For the approximation, assume that t_i are independent. Then the probability that no IAT exceeds I_T can be decomposed into the product of probabilities that each individual IAT does not exceed I_T, which is the cummulative distribution function of the $Beta(1, N)$ distribution, defined as $I_x(1, N) = 1 - (1 - x)^N$.

$$\begin{aligned}\mathbb{P}(\max\{t_1, ..., t_{N+1}\} \leq T_I) &= \mathbb{P}(t_1, ..., t_{N+1} \leq T_I) \\ &\approx \prod_{i=1}^{N+1} \mathbb{P}(t_i \leq T_I) \\ &= I_{\frac{T_I}{t}}(1, N)^{N+1} = \left(1 - \left(1 - \frac{T_i}{t}\right)^N\right)^{N+1}\end{aligned} \tag{32}$$

Figure 32 compares the approximation from Equation 32 with Monte Carlo simulation results from 3×10^5 random samples drawn from a flat Dirichlet distribution. It can be observed that the error term is comparably large for small N and gradually decreases for larger values of N, while taking a normal-distribution-like shape. Given that the approximation error is introduced by assuming independence of IATs, this can be explained by the law of large numbers. Specifically, with increasing N, the probability decreases that the sum of independently sampled IATs defers from its expectation (t), i.e., the IATs don't add up to t, introducing an error.

With an approximation of $\mathbb{P}(\max\{t_1, ..., t_{N+1}\} \leq T_I)$ obtained for a fixed N, the approximated probabilities can now be summed up for all $N = n$, weighted by the Poisson probability of $N = n$ packet arrivals during t.



(a) $N = 1$ (b) $N = 2$

(c) $N = 4$ (d) $N = 8$

Figure 32: Analytical Approximation vs. Monte-Carlo Simulation of $\mathbb{P}(\max\{t_1, ..., t_{N+1}\} \leq T_I)$ for $0 \leq \frac{T_I}{t} \leq 1$ and $N = \{1, 2, 4, 8\}$

$$\begin{aligned}\mathbb{P}(\text{“No timer expiry in time } t\text{”}) &= \sum_{n=1}^{\infty} \mathbb{P}(\mathcal{N}(t) = n)\mathbb{P}\big(\max\{t_1, ..., t_{N+1}\} \leq T_I \mid N = n\big) \\ &= \sum_{n=1}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} \left(1 - \left(1 - \frac{T_I}{t}\right)^n\right)^{n+1} = e^{-\lambda} \sum_{n=1}^{\infty} \frac{\lambda^n}{n!} \left(1 - \left(1 - \frac{T_I}{t}\right)^n\right)^{n+1}\end{aligned} \tag{33}$$

Although the summation includes an exponential series, the series cannot easily be brought into a closed form because it is weighted with another exponential term which depends on $n + 1$. Due to the lack of a closed form representation, it becomes evident that the resulting probabilities for the considered subproblem – despite already involving an approximation – are hard to compute, therefore defeating the purpose of an analytical solution.

List of Acronyms

3GPP **– 3rd Generation Partnership Project**

5QI **– 5G QoS Identifier**

BS **– Base Station**

BWP **– Bandwidth Part**

CDMA **– Code-Division Multiple Access**

C-DRX **– Connected Mode DRX**

CP **– Cyclic Prefix**

CSI **– Channel State Information**

CSI-RS **– Channel State Information Reference Signal**

DCI **– Downlink Control Information**

DL **– Downlink**

DRX **– Discontinuous Reception**

FDD **– Frequency-Division Duplexing**

FR **– Frequency Range**

HARQ **– Hybrid Automatic Repeat Request**

IAT **– Inter-Arrival Time**

i.i.d. **– independent and identically distributed**

LTE **– Long-Term Evolution**

MAC **– Medium Access Control**

MCS **– Modulation and Coding Scheme**

MIB **– Master Information Block**

MIMO **– Multiple Input, Multiple Output**

MTU **– Maximum Transmission Unit**

NR **– New Radio**

5G NSA **– 5G Non-Standalone**

OFDM **– Orthogonal Frequency-Division Multiplexing**

PBCH **– Physical Broadcast Channel**

PCB **– Printed Circuit Board**

PDCCH **– Physical Downlink Control Channel**

PDSCH **– Physical Downlink Shared Channel**

PoC **– Proof of Concept**

PPO **– Proximal Policy Optimization**

PRACH **– Physical Random Access Channel**

PU **– Power Unit**

PUCCH **– Physical Uplink Control Channel**

PUSCH **– Physical Uplink Shared Channel**
QAM **– Quadrature Amplitude Modulation**
QMI **– Qualcomm Mobile Station Modem Interface**
QoE **– Quality of Experience**
QoS **– Quality of Service**
QPSK **– Quadrature Phase Shift Keying**
RAN **– Radio Access Network**
RB **– Resource Block**
RE **– Resource Element**
RF **– Radio Frequency**
RIC **– RAN Intelligent Controller**
RL **– Reinforcement Learning**
RLC **– Radio Link Control**
RNTI **– Radio Network Temporary Identifier**
RRC **– Radio Resource Control**
RRM **– Radio Resource Management**
RSRP **– Reference Signal Received Power**
RTT **– Round-Trip Time**
Rx **– Receive**
5G SA **– 5G Standalone**
SCS **– Subcarrier Spacing**
SIB **– System Information Block**
SNR **– Signal-to-Noise Ratio**
SR **– Scheduling Request**
SRS **– Sounding Reference Signal**
SSB **– Signal Synchronization Block**
TCP **– Transmission Control Protocol**
TDD **– Time-Division Duplexing**
Tx **– Transmit**
UDP **– User Datagram Protocol**
UE **– User Equipment**
UL **– Uplink**
UML **– Unified Modeling Language**
USB **– Universal Serial Bus**

Bibliography

[1] D. H. Morais, "5G NR Overview," *5G NR, Wi-Fi 6, and Bluetooth LE 5: A Primer on Smartphone Wireless Technologies*. Springer Nature Switzerland, Cham, pp. 101–129, 2023. doi: 10.1007/978-3-031-33812-0_8.

[2] 3GPP, "5G; NR; Physical layer procedures for control," no. TS 38.213 v16.17.0. Apr. 2024.

[3] X. Lin, D. Yu, and H. Wiemann, "A Primer on Bandwidth Parts in 5G New Radio."

[4] A. Omri, M. Shaqfeh, A. Ali, and H. Alnuweiri, "Synchronization Procedure in 5G NR Systems," *IEEE Access*, vol. 7, pp. 41286–41295, 2019, doi: 10.1109/ACCESS.2019.2907970.

[5] 3GPP, "5G; NR; Radio Resource Control (RRC); Protocol specification," no. TS 38.331 v16.17.0. Aug. 2024.

[6] A. R. Jensen, M. Lauridsen, P. Mogensen, T. B. Sørensen, and P. Jensen, "LTE UE Power Consumption Model: For System Level Energy and Performance Optimization," in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, Sep. 2012, pp. 1–5. doi: 10.1109/VTCFall.2012.6399281.

[7] M. Towers *et al.*, "Gymnasium: A Standard Interface for Reinforcement Learning Environments." Accessed: Feb. 25, 2025. [Online]. Available: http://arxiv.org/abs/2407.17032

[8] S. Schneider, H. Karl, R. Khalili, and A. Hecker, "Multi-Agent Deep Reinforcement Learning for Coordinated Multipoint in Mobile Networks," *IEEE Transactions on Network and Service Management (TNSM)*, 2023.

[9] 3GPP, "5G; NR; Physical channels and modulation," no. TS 138.211 v16.10.0. Jul. 2022.

[10] 3GPP, "Study on User Equipment (UE) power saving in NR," Jun. 2019. [Online]. Available: https://atisorg.s3.amazonaws.com/archive/3gpp-documents/Rel16/ATIS.3GPP.38.840.V1600.pdf

[11] M. Lauridsen, D. Laselva, F. Frederiksen, and J. Kaikkonen, "5G New Radio User Equipment Power Modeling and Potential Energy Savings," in *2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall)*, Sep. 2019, pp. 1–6. doi: 10.1109/VTCFall.2019.8891215.

[12] U. Cisco, "Cisco annual internet report (2018–2023) white paper," *Cisco: San Jose, CA, USA*, vol. 10, no. 1, pp. 1–35, 2020, [Online]. Available: https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[13] A. Hassan *et al.*, "The Case for Boosting Mobile Application QoE via Smart Band Switching in 5G/xG Networks," in *Proceedings of the 25th International Workshop*

on Mobile Computing Systems and Applications, in HOTMOBILE '24. New York, NY, USA: Association for Computing Machinery, Feb. 2024, pp. 127–132. doi: 10.1145/3638550.3641132.

[14] H. Deng, Q. Li, J. Huang, and C. Peng, "iCellSpeed: increasing cellular data speed with device-assisted cell selection," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, in MobiCom '20. New York, NY, USA: Association for Computing Machinery, Sep. 2020, pp. 1–13. doi: 10.1145/3372224.3419201.

[15] 3GPP, "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; Vocabulary for 3GPP Specifications," no. TR 121 905 v16.1.0. Apr. 2024.

[16] T. Coughlin, "A Moore's law for mobile power," in *2015 IEEE Conference on Technologies for Sustainability (SusTech)*, Jul. 2015, pp. 137–140. doi: 10.1109/SusTech.2015.7314336.

[17] 3GPP, "5G; NR; User Equipment (UE) procedures in Idle mode and in RRC Inactive state," no. TS 138 304 v16.11.0. Jan. 2025.

[18] 3GPP, "5G; Procedures for the 5G System (5GS)," no. TS 23.502 v16.5.0. Jul. 2020.

[19] E. Coronado, S. Siddiqui, and R. Riggio, "Roadrunner: O-RAN-based Cell Selection in Beyond 5G Networks," in *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*, Apr. 2022, pp. 1–7. doi: 10.1109/NOMS54207.2022.9789832.

[20] V. Yajnanarayana, H. Rydén, and L. Hévizi, "5G Handover using Reinforcement Learning," in *2020 IEEE 3rd 5G World Forum (5GWF)*, Sep. 2020, pp. 349–354. doi: 10.1109/5GWF49715.2020.9221072.

[21] M. Mollel, S. Kaijage, and K. Michael, "Deep Reinforcement Learning based Handover Management for Millimeter Wave Communication," *International Journal of Advanced Computer Science and Applications*, 2021, doi: 10.14569/IJACSA.2021.0120298.

[22] M. Fiedler, T. Hossfeld, and P. Tran-Gia, "A generic quantitative relationship between quality of experience and quality of service," *IEEE Network*, vol. 24, no. 2, pp. 36–41, Mar. 2010, doi: 10.1109/MNET.2010.5430142.

[23] M. Fiedler and T. Hossfeld, "Quality of Experience-related differential equations and provisioning-delivery hysteresis," presented at the 21st ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE, IEICE, 2010. [Online]. Available: https://urn.kb.se/resolve?urn=urn:nbn:se:bth-7627

[24] S. Ickin, K. Wac, M. Fiedler, L. Janowski, J.-H. Hong, and A. K. Dey, "Factors influencing quality of experience of commonly used mobile applications," *IEEE*

Communications Magazine, vol. 50, no. 4, pp. 48–56, Apr. 2012, doi: 10.1109/MCOM.2012.6178833.

[25] M. S. Mollel *et al.*, "A Survey of Machine Learning Applications to Handover Management in 5G and Beyond," *IEEE Access*, vol. 9, pp. 45770–45802, 2021, doi: 10.1109/ACCESS.2021.3067503.

[26] E. Dahlman, S. Parkvall, and J. Skold, *5G NR: The Next Generation Wireless Access Technology*. Academic Press, 2020.

[27] T. Innovations, "LTE in a Nutshell," *White paper*, 2010, Accessed: Jul. 31, 2024. [Online]. Available: https://rintintin.colorado.edu/~gifford/5830-AWL/LTE%20in%20a%20Nutshell%20-%20Physical%20Layer.pdf

[28] R. A. K. Fezeu *et al.*, "Unveiling the 5G Mid-Band Landscape: From Network Deployment to Performance and Application QoE," in *Proceedings of the ACM SIGCOMM 2024 Conference*, in ACM SIGCOMM '24. New York, NY, USA: Association for Computing Machinery, Aug. 2024, pp. 358–372. doi: 10.1145/3651890.3672269.

[29] 3GPP, "5G; NR; Multiplexing and channel coding," no. TS 38.212 v16.14.0. Aug. 2024.

[30] I. Ahmed *et al.*, "A Survey on Hybrid Beamforming Techniques in 5G: Architecture and System Model Perspectives," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3060–3097, 2018, doi: 10.1109/COMST.2018.2843719.

[31] 3GPP, "5G; NR; Physical layer procedures for data," no. TS 38.214 v16.17.0. Aug. 2024.

[32] 3GPP, "5G; NR; Medium Access Control (MAC) protocol specification," no. TS 38.321 v16.16.0. Jul. 2024.

[33] P. F. Pérez, C. Fiandrino, and J. Widmer, "Characterizing and Modeling Mobile Networks User Traffic at Millisecond Level," in *Proceedings of the 17th ACM Workshop on Wireless Network Testbeds, Experimental evaluation & Characterization*, in WiNTECH '23. New York, NY, USA: Association for Computing Machinery, Oct. 2023, pp. 64–71. doi: 10.1145/3615453.3616509.

[34] 3GPP, "5G; System architecture for the 5G System (5GS)," no. TS 23.501 v16.20.0. Jul. 2024.

[35] R. S. Sutton and A. Barto, *Reinforcement learning: an introduction*, Second edition. in Adaptive computation and machine learning. Cambridge, Massachusetts London, England: The MIT Press, 2020.

[36] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, "Proximal Policy Optimization Algorithms." Accessed: Feb. 15, 2025. [Online]. Available: http://arxiv.org/abs/1707.06347

[37] J. Eschmann, "Reward Function Design in Reinforcement Learning," *Reinforcement Learning Algorithms: Analysis and Applications*. Springer International Publishing, Cham, pp. 25–33, 2021. doi: 10.1007/978-3-030-41188-6_3.

[38] Y. Yuan, Z. L. Yu, Z. Gu, X. Deng, and Y. Li, "A novel multi-step reinforcement learning method for solving reward hacking," *Applied Intelligence*, vol. 49, no. 8, pp. 2874–2888, Aug. 2019, doi: 10.1007/s10489-019-01417-4.

[39] S. Alraih, R. Nordin, A. Abu-Samah, I. Shayea, and N. F. Abdullah, "A Survey on Handover Optimization in Beyond 5G Mobile Networks: Challenges and Solutions," *IEEE Access*, vol. 11, pp. 59317–59345, 2023, doi: 10.1109/ACCESS.2023.3284905.

[40] 3GPP, "5G; NR; User Equipment (UE) radio access capabilities," no. TS 38.306 v16.19.0. Jan. 2025.

[41] V. Mnih *et al.*, "Human-level control through deep reinforcement learning.," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015, doi: 10.1038/nature14236.

[42] 3GPP, "5G; NR; NR and NG-RAN Overall description; Stage-2," no. TS 38.300 v16.18.0. Jan. 2025.

[43] S. Schneider, S. Werner, R. Khalili, A. Hecker, and H. Karl, "mobile-env: An Open Platform for Reinforcement Learning in Wireless Mobile Networks," in *NOMS 2022-2022 IEEE/IFIP Network Operations and Management Symposium*, Apr. 2022, pp. 1–3. doi: 10.1109/NOMS54207.2022.9789886.

[44] M. Lauridsen, L. Noël, T. B. Sørensen, and P. Mogensen, "An Empirical LTE Smartphone Power Model with a View to Energy Efficiency Evolution," *Intel Technology Journal*, vol. 18, no. 1, pp. 172–193, Mar. 2014.

[45] P. Wallentin, C. Tidestav, A. Orsino, and L. Qiu, "5G Advanced handover: L1/L2 Triggered mobility." Accessed: Sep. 03, 2024. [Online]. Available: https://www.ericsson.com/en/blog/2024/8/5g-advanced-handover-triggered-mobility

[46] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects," no. TR 36.814 v9.0.0. Mar. 2010.

[47] J. Navarro-Ortiz, P. Romero-Diaz, S. Sendra, P. Ameigeiras, J. J. Ramos-Munoz, and J. M. Lopez-Soler, "A Survey on 5G Usage Scenarios and Traffic Models," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 905–929, 2020, doi: 10.1109/COMST.2020.2971781.

[48] 3GPP, "5G; NR; Radio Link Control (RLC) protocol specification," no. TS 38.322 v16.3.0. Aug. 2022.

[49] "SIMXXXX-M2 4G/5G Hat Schematic." Accessed: Oct. 09, 2024. [Online]. Available: https://files.waveshare.com/upload/4/4d/SIMXXXX-M2_4G5G_HAT_SCH.pdf

[50] "SIM8200EA-M2 Hardware Design V1.03." Accessed: Oct. 09, 2024. [Online]. Available: https://files.waveshare.com/upload/d/de/SIM8200EA-M2_Hardware_Design_V1.03.pdf

[51] P. Jörke, H. Schippers, and C. Wietfeld, "Empirical Comparison of Power Consumption and Data Rates for 5G New Radio and RedCap Devices," presented at the 2025 IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas, USA, Jan. 2025. Accessed: Jan. 22, 2025. [Online]. Available: https://cni.etit.tu-dortmund.de/storages/cni-etit/r/Research/Publications/2025/Joerke_2025_CCNC/Joerke_CCNC2025_AuthorsVersion.pdf

[52] R. Jain, D. Chiu, and W. Hawe, "A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems." Accessed: Feb. 26, 2025. [Online]. Available: http://arxiv.org/abs/cs/9809099

[53] R. Durrett and R. Durrett, *Probability: Theory and Examples*. Cambridge University Press, 2019.

[54] K. Qi, Y. Chen, and W. Wu, "Dirichlet depths for point process," *Electronic Journal of Statistics*, vol. 15, no. 1, pp. 3574–3610, Jan. 2021, doi: 10.1214/21-EJS1867.