ALMA MATER STUDIORUM - UNIVERSITÀ DI BOLOGNA

SCHOOL OF ENGINEERING

DEPARTMENT OF ELECTRICAL, ELECTRONIC AND INFORMATION
ENGINEERING "GUGLIELMO MARCONI" - DEI

2nd Cycle Degree in
Telecommunications Engineering

Thesis on
Satellite Communications and Navigation Systems

A novel approach to Non-Orthogonal Random Access with application to NB-IoT NTN systems

Candidate: Supervisor:

Bruno De Filippo   Prof. **Alessandro Vanelli Coralli**

Cosupervisor:

Dr. **Carla Amatetti**

Academic Year 2021/2022

Session I

Contents

List of Figures

List of Tables

Introduction

In the massive IoT era, billions of low-complexity devices are expected to be connected to the Internet, posing challenges to the infrastructure and the technologies which these communications rely on. In particular, the improvement of the Random Access (RA) procedure to serve a higher amount of users has high priority, being this the first procedure that any device has to perform to connect to the network.

The objective of this thesis is to propose a novel approach to Non-Orthogonal Random Access (NORA), where multiple users can access the network using the same resources. This technology is even more important in Non-Terrestrial Networks, where cells can be several km wide and a massive amount of users have to be served in a short time window. Chapter 1 will present a brief introduction to the long range radio access standards for IoT with a focus on the chosen technology, Narrowband-IoT (NB-IoT). Then, Chapter 2 will analyze the steps of the RA procedure, detailing the rationale behind the RA preamble standardized in NB-IoT. This will be the ground for Chapter 3: after a State of the Art report on the NB-IoT RA procedure, along with the few published works on NORA, the proposed NORA technique will be presented. This includes the description of the slight modification to the RA preamble, which is required to ensure the detection of the preambles at the Base Station, and the definition of two detection algorithms for NORA. Chapter 4 will conclude this thesis, reporting the detection performance of the proposed algorithms. This evaluation will be carried out by means of simulation on MathWorks' computing environment MATLAB.

Chapter 1

Narrowband IoT

1.1 Wireless access technologies for IoT

The term *Internet of Things* (IoT) was first introduced by Kevin Ashton in 1999, when he proposed to implement *Radio Frequency Identification* (RFID) tags in P&G's supply chain to keep track of all of the resources utilized in real time, hence reducing wastes and losses [5]. Today, after more than 20 years, the IoT concept has evolved and spread over countless fields, adapting to different requirements and objectives. To give an idea of the order of magnitude of this technology, CISCO's latest Annual Internet Report for 2018-2023 registered more than 6 billions *Machine to Machine* (M2M) connections in 2018, while also estimating a growth to up to 14.7 billions in 2023 [6]. Moreover, a report by Gartner Inc. forecasted in August 2021 that the IT services for IoT market revenues will amount to $58 billion in 2025, up at 34% CAGR (*Compound Annual Growth Rate*) from 2020 [7]. It is clear that, even in the middle of a global chip shortage, the IoT concept is still expanding to more complex, wider networks. However, as the amount of data and devices increase, the wireless access technologies that realize these connections must remain up to the task. In the following paragraphs, the main wireless communication standards for IoT (for both licensed and unlicensed spectrum) will be introduced. Being the main part of this work based on *Narrowband IoT* (NB-IoT), an in-depth analysis of this technology will also be proposed in the second part of this chapter.

1.1.1 Licensed spectrum

3rd Generation Partnership Project (3GPP) has been working on standards for *Machine Type Communications* (MTC) scenarios since 2007 [8], leading to the development of M2M communication technologies in Release 13: *Extended Coverage Global System for Mobile Communications Internet of Things* (EC-GSM-IoT), *Long-Term Evolution for*

Coverage	164 dB
Throughput	160 bps
Latency	10 s
Network capacity	60'000 devices/km
Lifetime	10 years/5 Wh
Complexity	Ultra-low device complexity

Table 1.1: Objectives for EC-GSM-IoT, LTE-M e NB-IoT [3].

Machine-Type Communications (LTE-M) and *Narrowband Internet of Things* (NB-IoT). While the first two standards are based on already existing technologies (GSM and LTE, respectively), NB-IoT was developed from scratch, only taking into account some aspects of LTE to ensure some level of interoperability near already occupied frequency bands. At the foundation of these three standards is a set of targets (Table 1.1), which 3GPP considered mandatory to be satisfied by a competitive technology for *massive MTC* (mMTC) services. Due to the nature of these objectives, EC-GSM-IoT, LTE-M and NB-IoT are often called *Low-Power Wide Area Networks* (LPWANs).

The rapid expansion of the IoT market motivated 3GPP to include M2M-specific studies and features in the development of 5G since its start in 2015. Apart from the classic mobile communications objectives, expanded in the *enhanced Mobile Broadband* (eMBB) 5G use case, two new scenarios were defined: the aforementioned mMTC and *Ultra Reliable Low Latency Communications* (URLLC, also referred to as *critical MTC*, cMTC) [1]. Figure 1.1 illustrates how the three segments translate in real-life requirements on 5G, making this standard particularly versatile.

1.1.2 Unlicensed spectrum

3GPP's solutions are not the only IoT-focused wireless access technologies available on the market. Many organizations, like Sigfox and LoRa Alliance, developed standards to form LPWANs on bands that belong to the unlicensed portion of the frequency spectrum. It must be noted that the unlicensed bands are not necessarily the same across the globe, as the allocation is country dependent. As an example, *European Conference of Postal and Telecommunications Administrations* (CEPT) set the 863-870 MHz band as unlicensed in Europe [9], while in the United States of America the *Federal Communications Commission* (FCC) chose the 902-928 MHz band [10]. Furthermore, each country has restrictions on multiple electromagnetic emissions parameters, such as the channel utilization and the
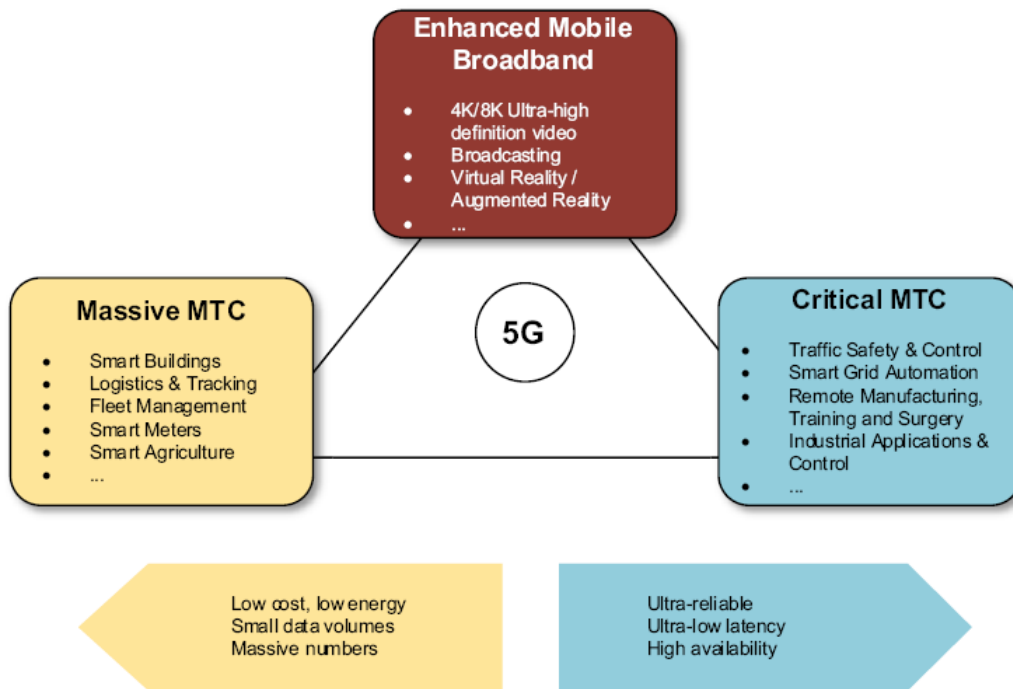
Figure 1.1: 5G segments and use cases [1].

maximum *Effective Isotropic Radiated Power* (EIRP). Finally, the same unlicensed band is shared between devices utilizing different communication protocols, meaning that some features must be introduced to ensure the coexistence of multiple technologies. It is easy to see that these issues make it far more difficult to implement a wide-area network with these wireless access technologies. The main wide-area standards on unlicensed spectrum are:

- *LoRa*, developed by the LoRa Alliance for the unlicensed bands below 1 GHz. It allows channels of 125 kHz in Europe and both 125 kHz and 500 kHz in the United States of America [11]. The standard utilizes *Chirp Spread-Spectrum* modulation, and it can reach data rates between 300 bps and 50 kbps depending on the particular service's requisites.

- *Ultra-Narrow Band* (UNB), developed by Sigfox, of which only a few characteristics are publicly available: as in LoRa, the devices are supposed to operate in the sub-1 GHz unlicensed band, but channels are 600 Hz large in the United States of America and 100 Hz large in the rest of the world [12], using a BPSK modulation in transmission. Data rates of some hundreds of bps can be reached by each connection.

- *Random Phase Multiple Access* (RPMA), developed by Ingenu to offer long range solutions for the 2.4 GHz bandwidth. As for UNB, only a small amount of information has been released to the public: the system utilizes a *Direct-Sequence*

Spread Spectrum (DSSS) modulation, with data rates around 5kbps for each MHz of occupied bandwidth [13].

1.2   NB-IoT

Among the technologies cited in Paragraph 1.1, NB-IoT is the IoT-dedicated solution with most advantages: EC-GSM-IoT and LTE-M are not subjected to the same level of flexibility, as it will be shown later on, while the standards on unlicensed spectrum are strongly limited by the coexistence of technologies that characterizes this portion of frequency spectrum. Finally, 5G is surely one of the most promising technologies, but it still lacks the maturity that the other technologies have. Studies on NB-IoT began in 2015, when LoRaWAN and Sigfox started to gain traction as alternatives to *Global System for Mobile Communications/General Packet Radio Service* (GSM/GPRS) for the LPWAN market. In order to develop a competitive wireless access technology, the study [3] was based on the set of requisites reported in Table 1.1 to ensure better performance than GSM/GPRS for IoT services.

One of the key characteristics of NB-IoT is its flexibility and compatibility with existing LTE networks. A NB-IoT carrier, which is 180 kHz wide, can be implemented in three ways:

- *Stand-alone* (Figure 1.2), mainly utilized for GSM spectrum refarming. Being GSM carriers 200 kHz wide and NB-IoT suggested guard bands 100 kHz wide, it can be concluded that two GSM carriers can be replaced by a single NB-IoT one.

- *In-band* (Figure 1.3, left), inside an LTE carrier using LTE frequency resources.

- *Guard-band* (Figure 1.3, right), occupying one of the two guard bands of an LTE carrier (these being large 5% of the carrier's bandwitdh).

The stand-alone deployment makes it easy to introduce one or more NB-IoT carriers inside a mobile operator's GSM spectrum, starting the spectrum refarming. The procedure can later be completed with a full migration from GSM to an LTE network: the existing stand-alone NB-IoT carriers will then become in-band or guard-band inside the new LTE carrier without any interruption of operation. Moreover, NB-IoT's flexibility allows for continuous operations even after a further migration from LTE to 5G, making this technology even more future-proof.
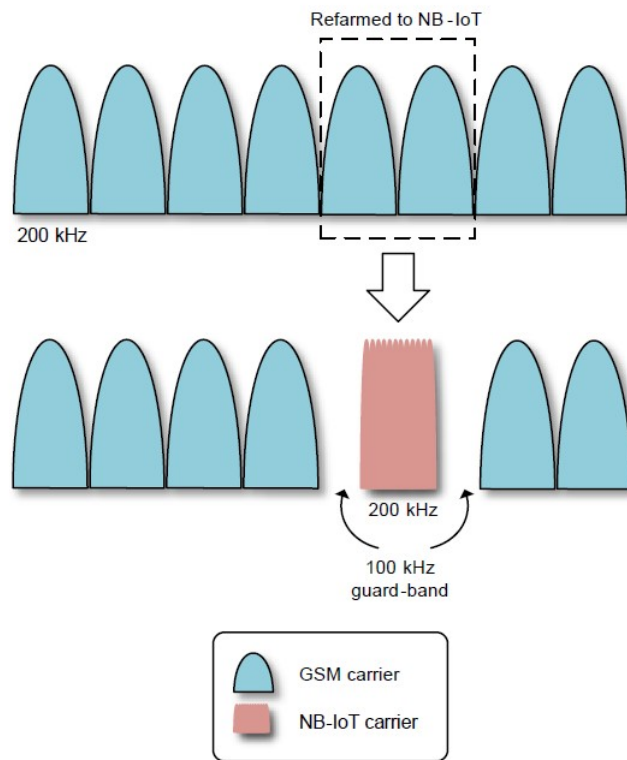
Figure 1.2: Stand-alone deployment in refarmed GSM spectrum [1].
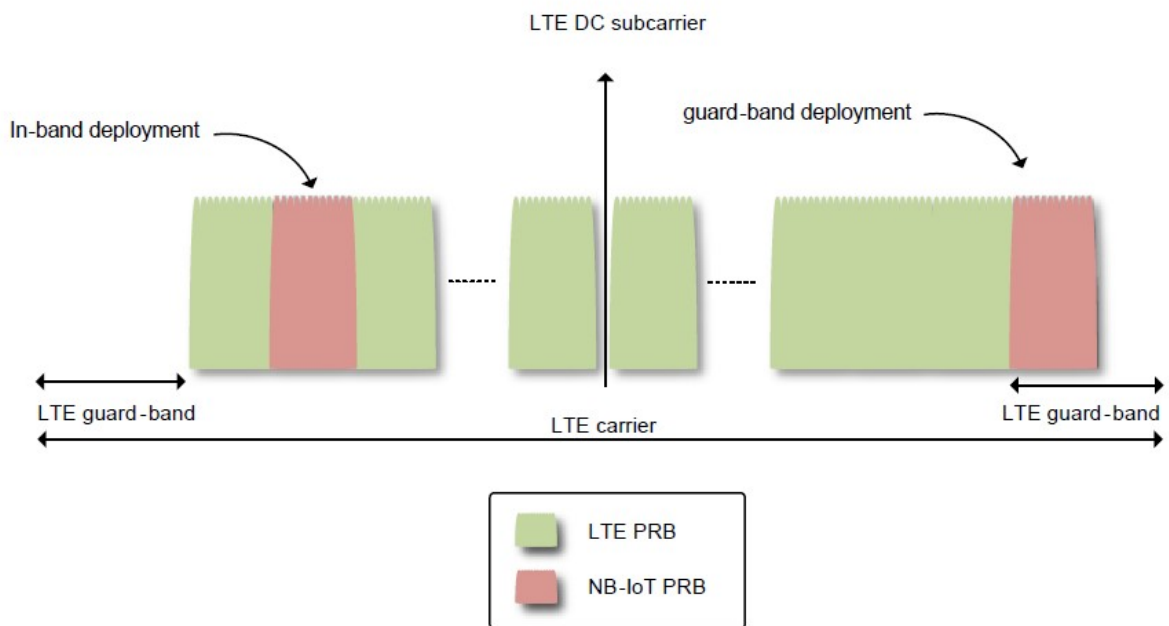
Figure 1.3: In-band (left) or guard-band (right) deployment in LTE spectrum [1].
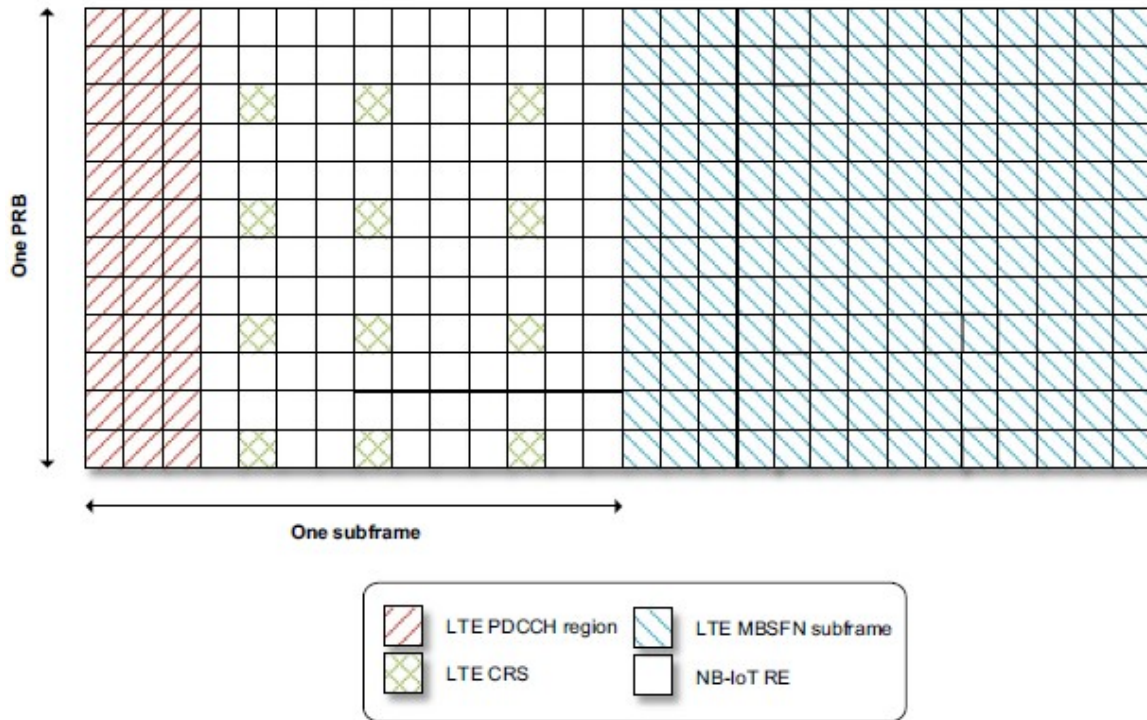
Figure 1.4: Downlink resource grid for NB-IoT, highlighting REs reserved to LTE [1].

1.2.1 Physical Layer

NB-IoT networks' coexistence with LTE is ensured by the physical layer design. In particular, NB-IoT utilizes the same waveforms of LTE, meaning that orthogonality is maintained and interference is reduced. Moreover, NB-IoT must use LTE's numerology, splitting time and frequency resources into a resource grid. Finally, in order to maintain interoperability with older LTE devices which may not be able to distinguish NB-IoT carriers from LTE ones, NB-IoT communications must happen in elements of the resource grid (called *Resource Elements*, REs) that are not reserved for essential LTE transmissions (e.g., synchronization signals). This last requisite strongly limits the resources fully available by NB-IoT users, as shown in Figure 1.4.

Following LTE's structure, the largest time period is the *hyperframe cycle*, composed by 1024 *hyperframes* each 10.24 s long. Each of these is made of 1024 *frames*, which are subdivided into 10 *subframes*. Finally, subframes are composed of 2 *slots*, which end up being 0.5 ms long (Figure 1.5). The subcarriers, which represent the frequency resources, are in general characterized by a 15 kHz spacing. However, while this structure has a general validity for both uplink (UL) and downlink (DL) communications, the UL resource grid structure can also be changed to smaller, 3.75 kHz-spaced subcarriers with frames directly divided into 5 slots (with a resulting duration of 2 ms each).
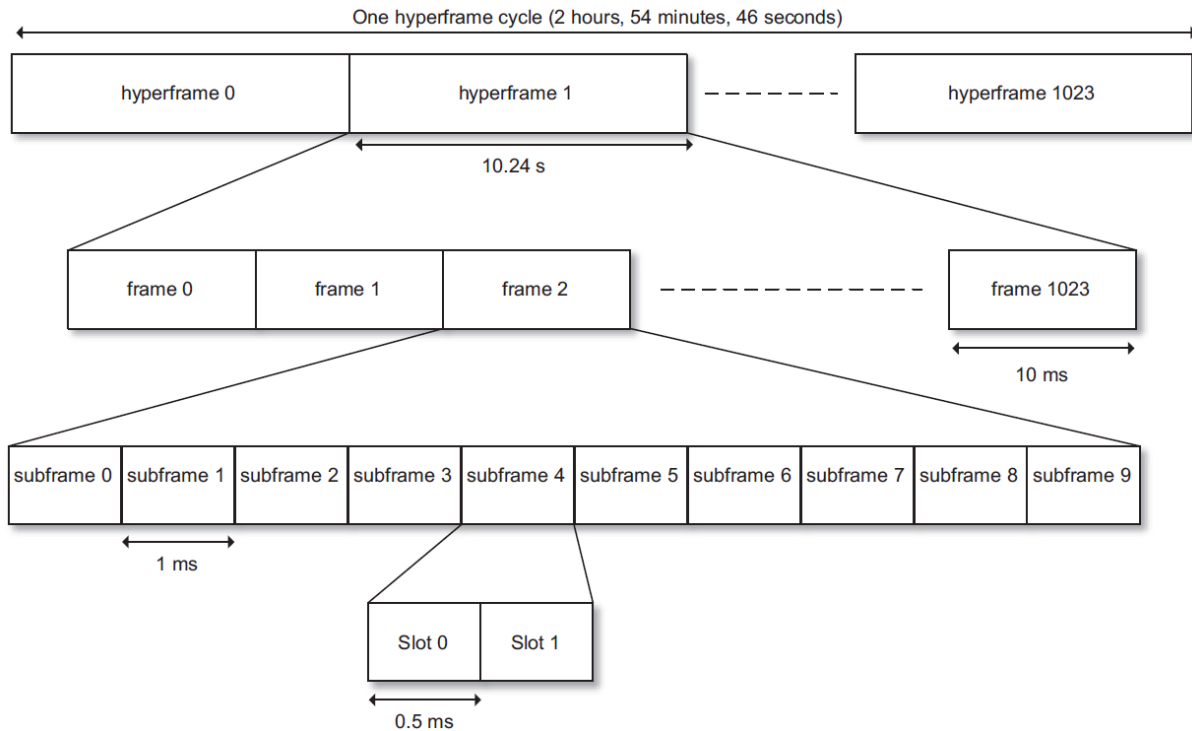
Figure 1.5: Time structure of NB-IoT for 15 kHz subcarrier spacing [1].

In order to specify the mapping of physical channels and signals onto REs, it is necessary to split the resource grid into time-frequency sections:

- In DL the concept of *Physical Resource Block* (PRB) is used, which corresponds to a portion of resource grid which spans 12 subcarriers and one slot. In most of the cases, a PRB pair (two time consecutive PRBs) is the smallest schedulable resource.

- In UL *Resource Units* (RUs) are implemented, coinciding with PRBs when 15 kHz subcarrier spacing is used, while in *single-tone allocation* (utilization of a single subcarrier with 3.75 kHz spacing) multiple formats are present.

DL transmissions utilize *Orthogonal Frequency-Division Multiple Access* (OFDMA) as in LTE, with a Cyclic Prefix (CP) 4.7 μs. The waveform is defined for 12 subcarriers and does not depend on the type of deployment. Being each time slot composed of 7 OFDM symbols, each PRB contains $12 \times 7 = 84$ REs. Following LTE specifications, NB-IoT's UL transmissions implement *Single-Carrier Frequency-Division Multiple Access* (SC-FDMA). Both 15 kHz and 3.75 kHz spaced subcarriers can be used individually (*single-tone transmission*) or multiple subcarriers with 15 kHz spacing can be aggregated (*multitone transmission*). In both cases, SC-FDMA is mathematically similar to an OFDM transmission scheme. Lastly, in order to improve coverage, NB-IoT exploits signal repetitions, which can be combined at the receiver in order to improve the SNR at the expense of a lower

data rate [14].

1.2.2   Physical channels

NB-IoT implements the following physical channels and signals for DL communications:

- *Narrowband Primary Synchronization Signal* (NPSS): contains time and frequency synchronization information.

- *Narrowband Secondary Synchronization Signal* (NSSS): provides further cell information.

- *Narrowband Reference Signal* (NRS): allows User Equipments (UEs) to estimate the channel's quality on the DL.

- *Narrowband Physical Broadcast Channel* (NPBCH): provides UEs with essential information to operate in the NB-IoT network.

- *Narrowband Physical Downlink Control Channel* (NPDCCH): provides UEs information about scheduled UL and DL resources, *paging* requests and base station (BS) information updates.

- *Narrowband Physical Downlink Shared Channel* (NPDSCH): utilized for both DL unicast and broadcast transmissions.

UL transmissions, on the other hand, require two channels:

- *Narrowband Physical Random Access Channel* (NPRACH): needed to initialize the connection of UEs to the network through the Random Access procedure.

- *Narrowband Physical Uplink Shared Channel* (NPUSCH): utilized for UL data, control information and acknowledgments (Acks) transfers.

In the next chapter, the NPRACH will be analyzed in depth.

Chapter 2

Random Access procedure in NB-IoT

Broadly speaking, the Random Access (RA) procedure is the operation carried out by UEs to request a connection setup from idle mode[15]. After a first synchronization phase, this procedure is the set of steps that a UE has to perform to be able to communicate in the network. The goals of the RA procedure are two-fold: first, to notify the network of the presence of a UE which wants to communicate; and second, to improve the UE synchronization with the chosen BS. Indeed, before the RA procedure, the UE synchronization is affected by the Time of Arrival (ToA) of the synchronization signals transmitted by the BS: with the RA procedure, each UE is able to compensate this to a large extent, hence improving the waveforms orthogonality and reducing interference. For the devices with lowest complexity (*NB-IoT Cat-N1*), the RA procedure is contention-based: this means that there are no mechanisms to ensure collisions cannot happen when multiple UEs try to access the channel. As a consequence, a dedicated step must be present to resolve these contention events.

The next paragraph will analyze in depth how the RA procedure works in a NB-IoT network.

2.1   RA procedure

The exchange of messages that composes the RA procedure is reported in Figure 2.1. In order to start the transmissions towards a BS, a UE must have correctly received all of the necessary synchronization information present in the synchronization signals reported in Chapter 1.2.2, while also ensuring that the access to the network is not barred (i.e., the UE is allowed to connect to the network). Among the received signals, NRS contains a sequence known to the UE, which allows the UE to perform channel estimation during the synchronization phase by measuring the *Reference Signal Received Power* (RSRP). Depending on this parameter and on the network settings, the UE determines which
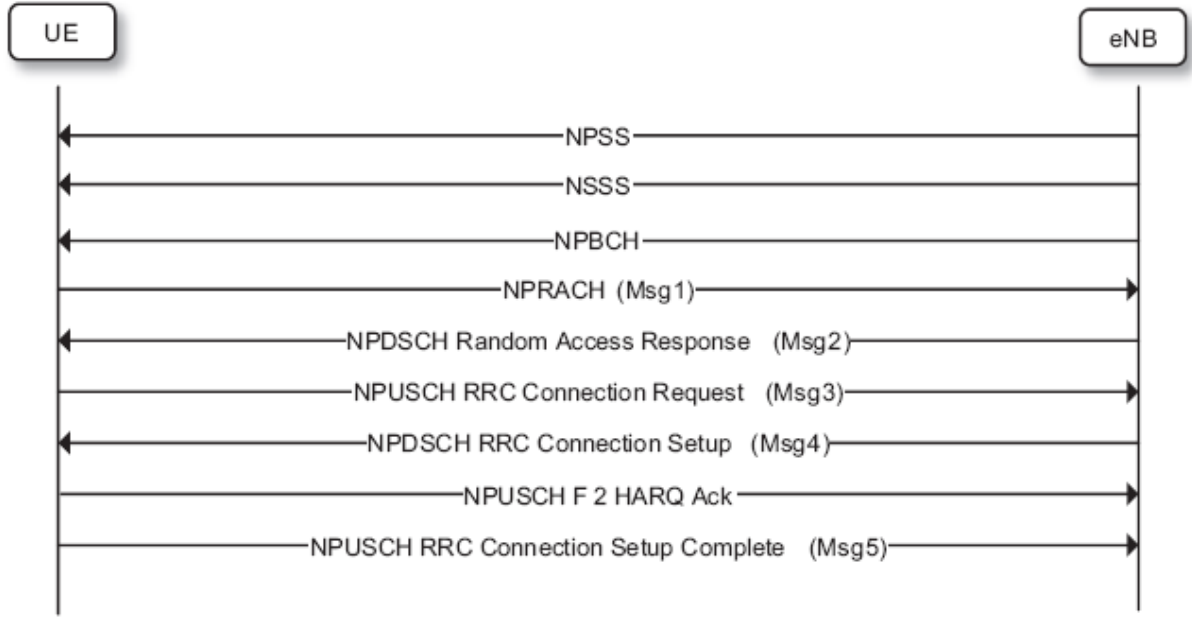
Figure 2.1: RA procedure steps in NB-IoT [1].

Number of NPRACH subcarriers	{12, 24, 36, 48}
Subcarrier offset	{0, 12, 24, 36, 2, 18, 34}
Periodicity	{40, 80, 160, 240, 320, 640, 1280, 2560} ms
Start time	{8, 16, 32, 64, 128, 256, 512, 1024} ms
Number of repetitions	{1, 2, 4, 8, 16, 32, 64, 128}

Table 2.1: NPRACH preamble parameters [4].

coverage class (CC) it belongs to. CCs are used in NB-IoT to allocate a different amount of resources to UEs based on their distance from the BS. Based on the CC (up to three can be implemented), UEs can start the transmission of the first RA message, named *NPRACH preamble*, only on specific time-frequency resources (Figure 2.2):

- The frequency domain allocation is defined by the number of subcarriers occupied by the NPRACH preamble and by the offset between the first NB-IoT subcarrier and the first NPRACH subcarrier.

- The time domain allocation is defined by the preamble's periodicity, the number of repetitions of the signal and a time offset between the start of the period and the start of the preamble. It must be noted that higher CCs (i.e., CCs for UEs further from the BS) will utilize a higher number of repetitions to improve the SNR.
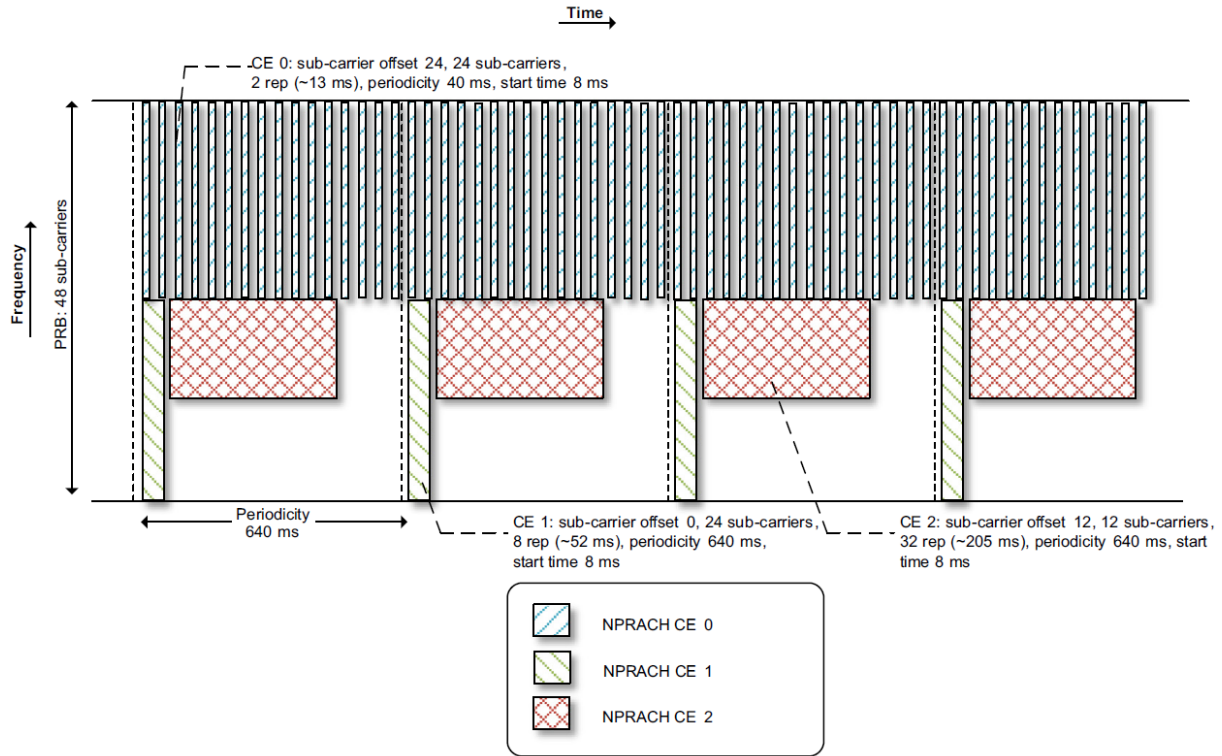
Figure 2.2: Example of NPRACH resources allocation [1].

Table 2.1 reports the values that the NPRACH preamble parameters can assume. The structure of the preamble will be further analyzed in the next sections. Once these information are obtained and the UE is synchronized with the BS, the transmission of the NPRACH preamble can happen. At the BS, an energy detector is implemented at the beginning of each RA opportunity: the received power is estimated over the resources that each possible NPRACH preamble can occupy; a preamble is then assumed to be present if the corresponding received power is higher than a properly fixed threshold. Knowing the signal's structure, the BS can compute a temporary identifier for each preamble detected, named *Random Access - Radio Network Temporary Identifier* (RA-RNTI), which depends on the frame number and the subcarrier on which the preamble started: this information will be used to specifically address the corresponding UE in the following messages. Furthermore, the BS can estimate the ToA of the received signal: from this, a parameter called *Timing Advance* (TA) can be computed, which represents the time offset that the UE must introduce to make its synchronization with the network as accurate as possible.

The BS can then transmit each TA to the correspondent UE through the *Random Access Response* (RAR). This message also contains the information needed by the UE to safely transmit in UL. To ensure that each device reads the correct RAR, each of these messages

is masked by the correspondent RA-RNTI.

Now, the UE can transmit a *Radio Resource Configuration* (RRC) *Connection Request* using the UL resources specified in the RAR: this message contains the UE (non-temporary) ID, a scheduling request and some additional information such as the receiver buffer status and the transmit power margin.

Thanks to the newly received ID, the BS can resolve any potential contention between users that chose the same RA preamble at the beginning of the procedure. Then, a *RRC Connection Setup* message is sent back to the UE to complete the connection setup.

Finally, the UE transmits its first message in connected mode, *RRC Connection Setup Complete*, which can already contain some data for the network. Before this, an Ack can possibly be sent for the last received message. This concludes the RA procedure in NB-IoT.

2.2   Design rationale of the NPRACH preamble

The structure of the NB-IoT's RA preamble is fundamentally different from the one of LTE's: this is mainly due to the fact that the waveforms utilized by the latter occupy 1 MHz of bandwidth, which is noticeably larger than NB-IoT's available bandwidth [1]. Furthermore, LTE's RA preamble is based on Zadoff-Chu sequences, which are characterized by a high *peak-to-average power ratio* (PAPR). This effect lowers the amplifiers efficiency and, as a consequence, the UEs' battery lifetime. To avoid these issues, the NPRACH preamble uses a single-tone transmission with frequency-hopping (FH). Three formats are available for the operator to choose from. The basic repetition unit of Format 0 and 1 is composed by 4 *symbol groups* (SGs), each containing a CP and 5 single-tone symbols on the same subcarrier (with spacing 3.75 kHz). The difference between these two formats lies in the CP length, which is 66.7 µs in Format 0 and 266.67 µs (which coincides with the symbol duration) in Format 1. In Format 2, on the other hand, the basic unit is composed of 6 SGs. Each of these contains a CP and 3 single-tone symbols 800 µs long on the same subcarrier, with a narrower spacing of 1.25 kHz (this value can be used only for NPRACH). This last format was introduced in Release 15 to improve the NPRACH preamble reception coverage, reaching up to 120 km [4]. It must be noted that, even if the terminology utilizes the word "symbol", the UE transmits just an unmodulated carrier during each SG.

In order to produce the signal, the UE selects a starting subcarrier between the ones in
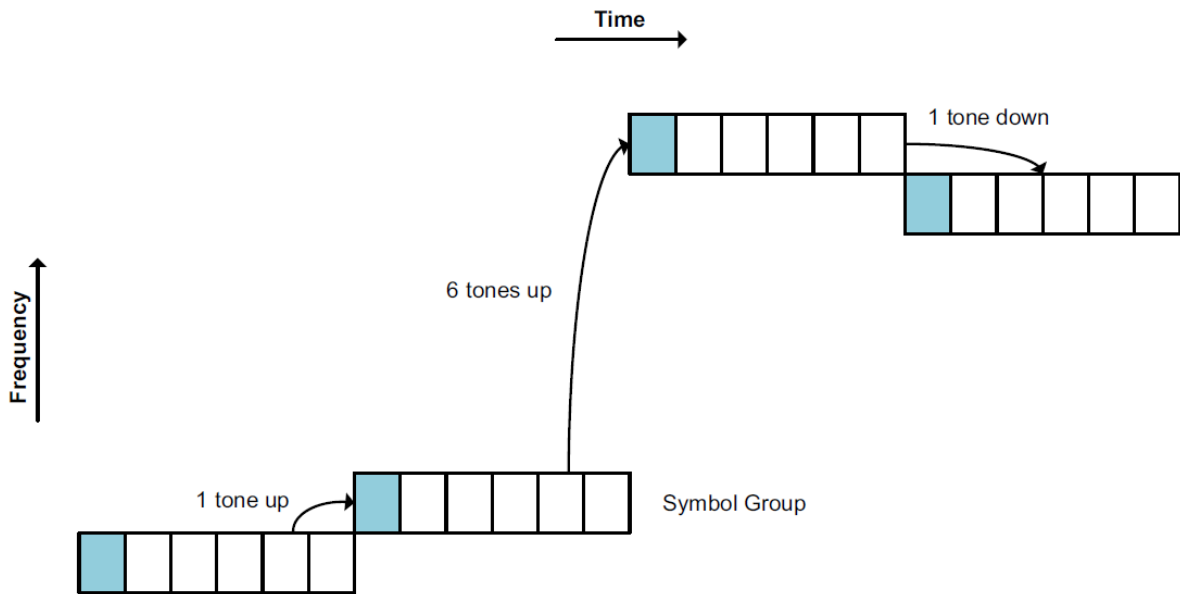
Figure 2.3: Example of time-frequency resources utilization of a NPRACH preamble, Format 1 [1].

the NPRACH band: this will be the tone used to transmit the first SG. In Format 0 and 1, the set of subcarriers over which the remaining SGs are transmitted is determined from the starting subcarrier (Figure 2.3). In particular, every pattern requires a hop by one subcarrier between the first and second SGs and between the third and fourth SGs, while between the second and third SGs the hop is 6 subcarriers long. Overall, the 4 SGs long repetition unit will lie in a 12 subcarriers frequency range. Given that the NPRACH band can have up to 48 subcarriers and that the patterns are deterministic, only 48 patterns can be present at most. This limit can strongly delay the connection of a large amount of users, which is the main reason that prompted this work: in Chapter 3 a solution to noticeably reduce this delay will be proposed.

In Format 2, a 36 subcarriers bandwidth is available for the frequency-hopping pattern, thanks to the 1.25 kHz subcarrier spacing. In this format, the deterministic patterns are composed by a hop by a single subcarrier between the first and the second SGs and between the the fifth and the sixth SGs, 3 subcarriers between the second and the third SGs and between the fourth and the fifth SGs and finally an 18 subcarriers long hop between the third and the fourth SGs. The rationale behind this pattern is that the abrupt change in frequency between SGs allows the BS to better determine the ToA of the signal, hence improving the TA that will be sent to the UE in the following step of the RA procedure. The parameters that characterize the three formats are summarized in Table 2.2.

	Format 0	Format 1	Format 2
Supported cell radius	10 km	40 km	120 km
CP length	66.67 µs	266.67 µs	800 µs
Symbol length	266.67 µs	266.67 µs	800 µs
Symbols per SG (no CP)	5	5	3
SGs per repetition unit	4	4	6
Subcarrier spacing	3.75 kHz	3.75 kHz	1.25 kHz
Frequency-hopping range	12 subcarriers	12 subcarriers	36 subcarriers

Table 2.2: NPRACH preamble allowed parameters for Format 0, 1 and 2 [4].

However, in order to reach a satisfactory SNR level, one transmission could not be enough. Hence, the basic unit can be repeated for a maximum of 128 times: in this way, noise can be averaged out and even UEs far from the BS can be correctly detected. To reduce the risk of persistent interference among users that are performing the RA procedure with the same preamble in two distinct cells, a second hopping pattern, named *pseudo-random hopping pattern*, maps each basic unit to a different set of 12 (or 36 for Format 2) subcarriers. This pattern is computed as a function of the cell ID and the repetition index, such that two NPRACH preambles intended for two distinct BSs do not utilize the same set of subcarriers for the entire duration of the transmission. It must be noted that if two users inside the same cell choose the same NPRACH preamble (i.e., choose the same starting subcarrier), the pseudo-random hopping pattern will not help to avoid the full collision. Nonetheless, this contention will be managed by the BS in the next steps of the RA procedure as reported in Section 2.1.

Chapter 3

From Orthogonal Random Access to NORA

In order to present the proposed modifications to the NPRACH preamble that can enable NORA in NB-IoT, it is necessary to analyze in-depth the transmitted waveform in the traditional Orthogonal Random Access (ORA) technique.

3.1   State of the Art of ORA

With the design considerations from Section 2.2 in mind, the i-th transmitted symbol of the m-th SG can be mathematically described as follows:

$$s_{m,i}(t) = \sqrt{P_T} \times e^{j2\pi \Delta f f_k(m)t},\tag{3.1}$$

where $f_k(m)$ is the subcarrier index over which the m-th Symbol Group (SG) is transmitted following the k-th FH pattern, Δf is the subcarrier spacing (3.75 kHz) and P_T is the signal power. An example of a generated NPRACH preamble with 2 repetitions is reported in Figure 3.1.

In order to determine the subcarrier index utilized for the transmission of each SG, an algorithm runs at the UE. Being:

- N_{sc}^{NPRACH} the number of available subcarriers for NPRACH, chosen between {12, 24, 36, 48};

- N_{offset}^{NPRACH} the offset between the first NPRACH subcarrier and the first subcarrier of the considered PRB;

- $N_{sc}^{RA} = 12$ the number of available subcarriers for the FH pattern (Format 1);
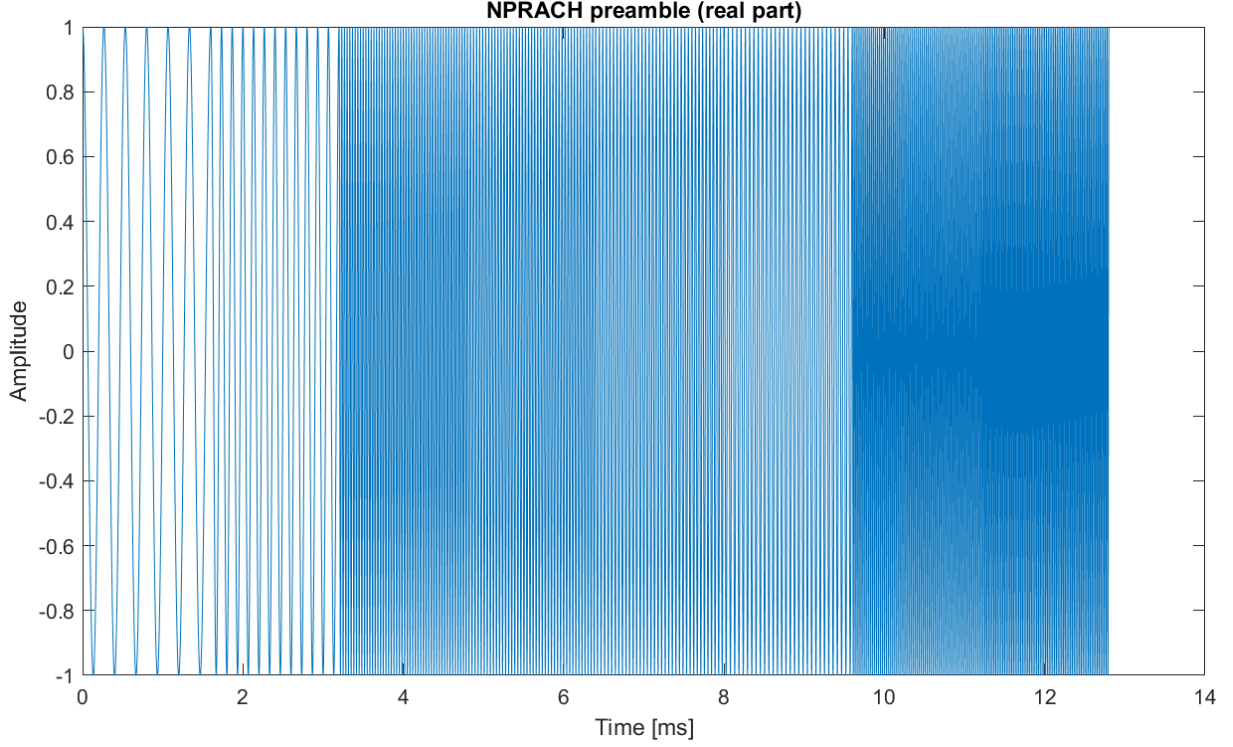
Figure 3.1: Example of transmitted NPRACH preamble (real part), Format 1.

- n_{init} the subcarrier index over which the first SG is transmitted, relative to the NPRACH band. This value is chosen randomly from a uniform distribution of the integer values between 0 and $N_{sc}^{NPRACH} - 1$;

- $n_{start} = N_{offset}^{NPRACH} + \left\lceil \frac{n_{init}}{N_{sc}^{RA}} \right\rceil \times N_{sc}^{RA}$ the subcarrier index over which the first SG is transmitted, relative to the entire NB-IoT band;

the subcarrier over which the m-th SG is transmitted is $n_{sc}^{RA}(m) = n_{start} + \tilde{n}_{sc}^{RA}(m)$, where $\tilde{n}_{sc}^{RA}(m)$ is determined through the following set of equations:

$$\tilde{n}_{sc}^{RA}(0) = n_{init} \ mod \ N_{sc}^{RA}$$

$$\tilde{n}_{sc}^{RA}(m) = \begin{cases} \left(\tilde{n}_{sc}^{RA}(0) + f(m/4)\right) mod \ N_{sc}^{RA} & \text{if } m \ mod \ 4 = 0 \text{ and } m > 0 \\ \tilde{n}_{sc}^{RA}(m-1) + 1 & \text{if } m \ mod \ 4 = 1, 3 \text{ and } \tilde{n}_{sc}^{RA}(m-1) \ mod \ 2 = 0 \\ \tilde{n}_{sc}^{RA}(m-1) - 1 & \text{if } m \ mod \ 4 = 1, 3 \text{ and } \tilde{n}_{sc}^{RA}(m-1) \ mod \ 2 = 1 \\ \tilde{n}_{sc}^{RA}(m-1) + 6 & \text{if } m \ mod \ 4 = 2 \text{ and } \tilde{n}_{sc}^{RA}(m-1) < 6 \\ \tilde{n}_{sc}^{RA}(m-1) - 6 & \text{if } m \ mod \ 4 = 2 \text{ and } \tilde{n}_{sc}^{RA}(m-1) \geq 6 \end{cases}$$

$$(3.2)$$

$$f(t) = \left(f(t-1) + \left(\sum_{n=10t+1}^{10t+9} c(n) 2^{n-(10t+1)} \right) mod \ \left(N_{sc}^{RA} - 1\right) + 1 \right) mod \ N_{sc}^{RA}$$

$$f(-1) = 0.$$

Here, $c(n)$ is a Gold sequence of length 31 which depends on N_{ID}^{cell}, the cell identifier [16]: this is the component that generates the pseudo-random FH pattern. As previously discussed, Equation 3.2 shows that $\tilde{n}_{sc}^{RA}(m)$ exclusively depends on the subcarrier chosen to transmit the first SG and on the pseudo-random FH pattern $f(t)$.

The NPRACH preamble has been widely investigated in the literature. [17] reported the waveform's main characteristics, explaining its design rationale and proposing both a preamble detection and ToA estimation algorithms. The paper provided the performance of these algorithm assuming an Additive White Gaussian Noise (AWGN) channel. [16] followed with the analysis of the NPRACH preamble detection performance reporting the Correct Detection (CD) and False Alarm (FA) probability estimates while varying the threshold value. The results of these two studies were improved in [18], where the detection of NPRACH preambles under a Rayleigh fading channel was investigated; also, the synchronization parameters estimation was studied in this paper. Finally, in [19] the performance of a receiver over the uplink NB-IoT channels (NPRACH included) was obtained, presenting a physical implementation on a commercial LTE BS. These results were then compared with the standard's requirements to ensure the system's full compliance. Focusing more on NTN-specific researches, a RA receiver was proposed in [20] to reduce the performance degradation due to power leakage of NPRACH preambles over adjacent subcarriers. In particular, the solution proposed by the authors coherently combines the adjacent subcarriers to take advantage of the leaked power, while also trying to suppress interference from other users. In [21], a Wavelet Transform-based algorithm was investigated to detect the NPRACH preamble and estimate its ToA without the need to compensate for the Carrier Frequency Offset (CFO) first. Finally, in [2] a RA technique robust to typical NTN channel impairments (long delays, significant Doppler effects, wide beams) without the need to modify the preamble waveform. Here, a GNSS-free partial compensation of both the delay and the Doppler shift was also presented. In the mMTC landscape, where the UE complexity should be kept at a minimum, the on-board GNSS module can have low accuracy (leading to imprecise delay and Doppler estimates) or even be non-existent. However, while each UE is subject to different levels of impairments, the beam geometry in the NTN environment makes easy to split both the delay and Doppler shift into a common component, which each UE is subject to, and a UE-specific differential component as shown in Figure 3.2. The two common components can then be computed and pre-compensated at the BS by knowing the satellite's ephemeris, the beam center and the local topography, without any computation needed at the UE side. This leaves only the differential components as the main NTN channel impairments instead of the (much larger) entire delay and Doppler shift.
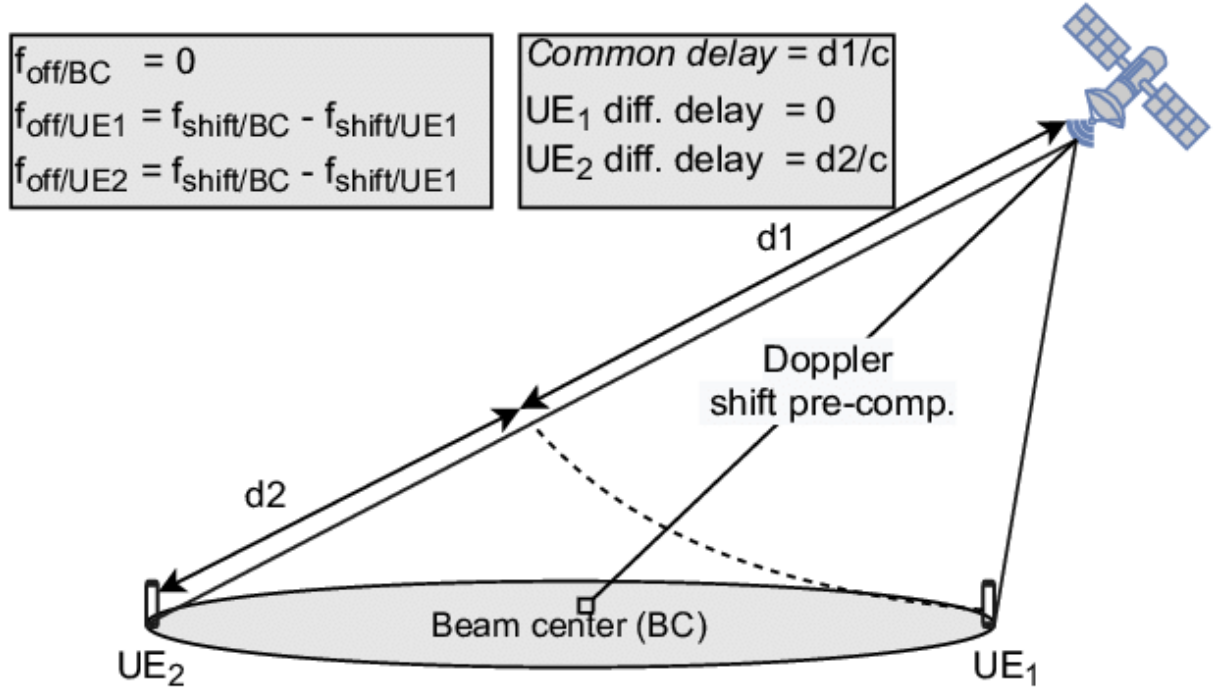
$f_{off/BC} = 0$
$f_{off/UE1} = f_{shift/BC} - f_{shift/UE1}$
$f_{off/UE2} = f_{shift/BC} - f_{shift/UE1}$

Common delay = d1/c
UE_1 diff. delay = 0
UE_2 diff. delay = d2/c

d1

Doppler
shift pre-comp.

d2

Beam center (BC)

UE_2 UE_1

Figure 3.2: Differential delay and Doppler [2].

3.2   Problem statement

The contention-based nature of the RA procedure suggests that multiple attempts could
be needed for a UE to successfully connect to the network. As stated in Section 2.2, in NB-
IoT a maximum of 48 distinct UEs can be detected every RA opportunity. However, NTNs
are characterized by large coverage areas and a very short coverage time: depending on
the satellite height, minimum elevation angle and UE location, Low Earth Orbit satellites
can remain in visibility for just a few minutes [22] and cells can be thousands of kilometers
wide. In the envisioned mMTC scenarios, this means that thousands of UEs could try
to access the network at the same time, introducing an excessive amount of collisions.
This would lead to severe bottlenecks and congestion, as the contended UEs would try to
connect again to the network during the next RA opportunities, introducing even more
collisions with new UEs. As a consequence, the system-level RA success probability would
rapidly decrease, leading to an extremely high number of retransmissions and, even more
importantly, unsustainable UE energy consumptions.

It is clear that there is a need to provide a rapid connection for an extremely large number
of UEs. Moreover, being NB-IoT able to fully coexist with LTE and 5G, its resources are
limited to the small amount of Resource Elements (REs, time-frequency resource unit)
that are not essential to LTE and 5G transmissions: indeed, LTE and 5G UEs must be
able to receive the frequent synchronization and control signals without any interference

from a coexistent NB-IoT network. This implies that the RA issue cannot be solved by allocating more resources to NB-IoT. This prompted the research on ways to improve the NB-IoT RA procedure, i.e., increasing the number of detectable users in the same time period, leading to the study of non-orthogonal techniques for RA. Following the 5G terminology for Non-Orthogonal Multiple Access (NOMA), this procedure is sometimes called Non-Orthogonal Random Access (NORA).

3.3   State of the Art of NORA

The concept of NORA has not been widely investigated in the literature. First, in [23] a Sparse Code Multiple Access (SCMA)-based technique was proposed for LTE/LTE-Advanced and 5G (where the RA procedure's structure is the same as reported in Section 2.1) to enable large-scale RA for mMTC. Here, Message 3 is transmitted using a De-Modulation Reference Signal (DMRS) out of the SCMA codebook to perform code domain spreading: this allows the reception of multiple messages, as long as they are sent using different DMRSs. Then, [24] investigated an unsupervised Deep Learning (DP)-based power control scheme based only on the TA information contained in the RAR. In [25], a combination of two SIC algorithms was proposed in order to improve the detection of superimposed Message 3s. However, all of these solutions only focused on the RRC Connection Request (Message 3), trying to resolve the contentions instead of reducing them.

On the opposite, [26] presented an overload control method able to separate multiple preambles by computing the power delay spectrum distribution of the received signal at the BS. By sending power control commands to colliding UEs, the contention can be resolved by performing SIC on Message 3. Nevertheless, the authors specify that their solution currently (as of April 2021) supports collisions of only two preambles at most: it is clear that these preliminary results, although promising, are not enough for the mMTC scenario.

Finally, the only NB-IoT-related NORA article in the literature is [27], where a Power Delay Profile (PDP) analysis (similar to that implemented in [26]) is performed to separate multiple NPRACH preambles. Here, the RARs destined to collided UEs are multiplexed through NOMA, meaning that each UE is supposed to perform SIC on Message 2 to obtain its own RAR. This allows the system to solve the contention earlier, but puts a strain on the UEs, which is not advisable in mMTC scenarios. Moreover, this solution doesn't provide a way to individually address collided UEs: indeed, the authors suggest that each UE should estimate its own TA, based on the received power over a synchronization signal, and determine its RAR by choosing the one which reports the TA parameter closest

to the estimated value.

It is clear that there is a need to improve the number of distinct NPRACH preambles while maintaining the UEs addressability and not increasing the resources occupied. In the remaining part of this chapter, a solution to this demand will be proposed: the Silenced Tone pattern.

3.4   The Silenced Tone pattern

In order to distinguish multiple preambles and maintain the ability to address each UE individually, a slight modification of the preamble structure is here proposed. Starting from the standard NPRACH preamble waveform described in Equation 3.2, each UE randomly chooses one of the SGs over which its transmitter will be turned off. In this way, two UEs can choose the same preamble but turn off their transmitter during two distinct time intervals: at the BS, the two signals will only partially collide, making the detection of both the preambles possible. This procedure introduces what can be seen as a Code Division Multiple Access (CDMA) technique in which each preamble is modulated by a low frequency code, named *Silenced Tone* (ST) *pattern*, of $N_{SG} - 1 = N_{rep} \times 4 - 1$ ones and a single zero. Mathematically, the i-th symbol of the m-th SG of a generic transmitted preamble can be expressed as follows:

$$s_{m,i}(t) = a_h(m) \times \sqrt{P_T} \times e^{j2\pi\Delta f f_k(m)t}, \tag{3.3}$$

$$a_h(m) = \begin{cases} 0 \text{ if } m = h \\ 1 \text{otherwise} \end{cases}$$

where the couple (h,k) represents the index of the chosen ST and FH patterns, respectively, and $a_h(m)$ implements the ST pattern (i.e., it sets the preamble to 0 when the SG index is equal to the ST pattern index). Figure 3.3 reports an example of a generated waveform where the transmitter was turned off during the second SG.
 It must be noted that, due to limitations introduced by the standard preamble detection at the BS, the first SG cannot be silenced. Nonetheless, the proposed solution is able to extend the number of available preambles by a factor $N_{SG} - 1$, thus potentially reducing the number of collisions.

As multiple UEs might be performing the RA procedure, the BS receives the sum of all of the transmitted preambles. In general, some UEs may choose the same FH and ST pattern: from now on, the number of users that transmitted a preamble with pattern (h,k)
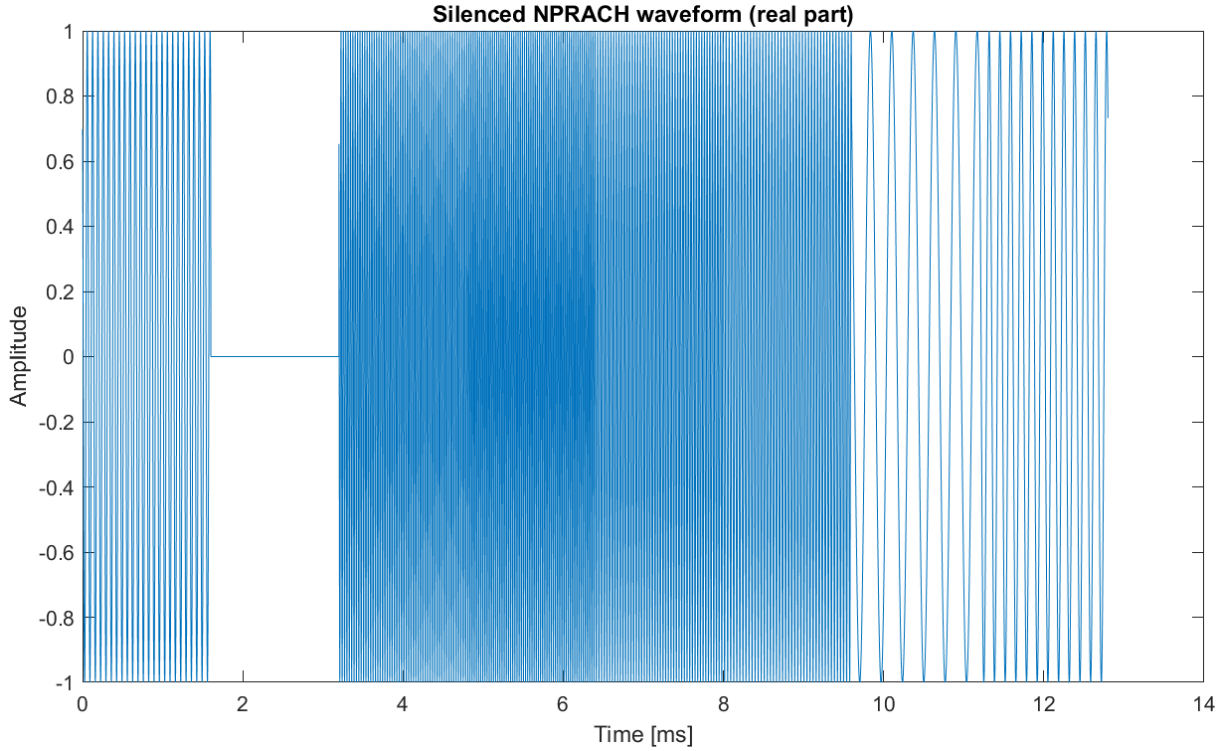
Figure 3.3: Example of transmitted silenced NPRACH preamble (real part), Format 1.

will be referred to as $N_{UE_{h,k}}$. Each signal will be individually phase-shifted by an amount which depends on the actual position of the UE, the satellite and the BS. This phase shift will be here considered as a random sample from a uniform distribution between 0 and 2π, assuming that the differential propagation delay left from the pre-compensation presented in Section 3.1 is lower than the CP length. Considering an ideal Additive White Gaussian Noise (AWGN) channel, the signal at the receiver can then be expressed as the sum of the superimposed preambles and Gaussian noise. Focusing on the generic K-th FH pattern (as FH ensures orthogonality between preambles using distinct FH patterns), the i-th received symbol of the m-th SG can be written as:

$$r_{m,i}(t) = \sum_{h=1}^{N_{SG}-1} \sum_{u=0}^{N_{UE_{h,K}}-1} a_h(m) \times \sqrt{P_{R,u}(h,K)} \times e^{j(2\pi\Delta f \times f_K(m)t+\theta_u(h,K))} + w(t), \quad (3.4)$$

where, considering the u-th user that transmitted the preamble with pattern (h,K), $P_{R,u}(h,K)$ is its received power, $\theta_u(h,K) \sim U[0,2\pi]$ is its random phase shift due to propagation and $w(t)$ represents a sample of a bandlimited Gaussian noise random process with variance $\sigma_n^2 = N_0/2$ (i.e., the AWGN bilateral power spectral density).

The BS can then sample the signal to obtain its discrete time representation. Assuming the BS uses the LTE-compliant sample rate $f_s = 1.92MHz$ and the NPRACH preambles

are generated with Format 1 numerology (i.e., symbol time $T_s = 266\mu s$), each symbol is represented by $L = f_s \times T_s = 512$ samples. Hence, each received symbol is sampled at instants $t_n = n\frac{T_s}{L}$, $n = 0, ..., L - 1$, obtaining the following discrete time signal:

$$r_{m,i}(n) = \sum_{h=1}^{N_{SG}-1} \sum_{u=0}^{N_{UE_{h,K}}-1} a_h(m) \times \sqrt{P_{R,u}(h, K)} \times e^{j(2\pi \times f_K(m)\frac{n}{L} + \theta_u(h,K))} + w(n). \quad (3.5)$$

It can be noticed that the subcarrier spacing Δf is now absent from Equation 3.5: this is due to the fact that it gets simplified with its reciprocal, the symbol time $T_s = \frac{1}{\Delta f}$, when it is introduced in the formula during sampling.

3.5 Energy Detector in NORA

At the BS, an Energy Detector can be implemented to detect NPRACH preambles [18]. However, the algorithm relies on the received power, which in NORA can vary significantly depending on the level of interference between non-orthogonal preambles. Hence, a statistical characterization of the received power should be carried out to determine whether an energy-based detector can be implemented in a NORA receiver.

The first operations that a BS performs on the baseband preamble are the CP removal and the application of the L-point Fast Fourier Transform (FFT) on each received symbol. The $f_K(m)$-th FFT bin of the i-th symbol of the m-th SG will then be:

$$R(m, i, f_K(m)) = \sum_{h=1}^{N_{SG}-1} \sum_{u=0}^{N_{UE_{h,K}}-1} a_h(m) \times \sqrt{P_{R,u}(h, K)} \times e^{j\theta_u(h,K)} + W_K(m, i), \quad (3.6)$$

where $W_K(m, i)$ is the post-FFT AWGN sample with variance $L \times \frac{N_0}{2}$. Hence, the FFT introduces an amplification effect on the noise power, which becomes $P_n = L \times N_0$.

While the current detection is performed over the entire preamble duration, the proposed NORA technique requires a decision on a SG basis. The average received power over the M-th SG of the K-th FH pattern can be computed by collecting the energy per symbol over the said SG and dividing it by the number of symbols per SG, which after the CP removal is 5:

$$P_{M,K} = \frac{1}{5} \sum_{i=0}^{4} |R(M, i, f_K(M))|^2 =$$

$$= \frac{1}{5} \sum_{i=0}^{4} \left| \sum_{h=1}^{N_{SG}-1} \sum_{u=0}^{N_{UE_{h,K}}-1} a_h(M) \times \sqrt{P_{R,u}(h, K)} \times e^{j\theta_u(h,K)} + W_K(M, i) \right|^2. \quad (3.7)$$

Following the mathematical derivation of [18], the statistical distribution of $P_{M,K}$ when both the signal and noise are present is a scaled non-central Chi-square with 10 degrees of freedom. This is due to fact that $P_{M,K}$ is the average of 5 independent complex Gaussian random variables (RVs), i.e., the average of 10 independent real Gaussian RVs. However, the distribution parameters are still to be evaluated:

- The scale factor k can be computed as $k = \frac{10}{P_n}$;

- The non-centrality parameter γ_0 is given by $\gamma_0 = 10 \frac{P_{0_{M,K}}}{P_n}$, where $P_{0_{M,K}}$ is the average received power over the M-th SG of the K-th FH pattern in absence of noise.

To compute $P_{0_{M,K}}$, the same procedure introduced for $P_{M,K}$ can be used, i.e., collect energy over the M-th SG and divide by 5. However, in absence of noise, the post-FFT bins are independent of the symbol index i, meaning that the following relationship can be obtained:

$$
\begin{aligned}
P_{0_{M,K}} &= \frac{1}{5} \sum_{i=0}^{4} \left| \sum_{h=1}^{N_{SG}-1} \sum_{u=0}^{N_{UE_{h,K}}-1} a_h(M) \times \sqrt{P_{R,u}(h,K)} \times e^{j\theta_u(h,K)} \right|^2 \\
&= \left| \sum_{h=1}^{N_{SG}-1} \sum_{u=0}^{N_{UE_{h,K}}-1} a_h(M) \times \sqrt{P_{R,u}(h,K)} \times e^{j\theta_u(h,K)} \right|^2 .
\end{aligned}
\tag{3.8}
$$

In order to proceed with a mathematical derivation of $P_{0_{M,K}}$, it is necessary to introduce a temporary assumption. From now on, the problem will be simplified by assuming that at most one UE can choose the preamble (h,k), i.e., $N_{UE_{h,k}} \in \{1,0\}$. The obtained result will then be generalized to $N_{UE_{h,k}} \in \mathbb{N}$. Naming N_{UE_k} the set of ST patterns used for the transmission of a preamble, i.e.,

$$
N_{UE_k} = \{h \text{ s.t. } N_{UE_{h,k}} = 1\},
\tag{3.9}
$$

and recalling the definition of $a_h(M)$ in Equation 3.3, the average power over the M-th SG of the K-TH FH pattern becomes:

$$
P_{0_{M,K}} = \left| \sum_{h \in N_{UE_K} \setminus \{M\}} \sqrt{P_R(h,K)} \times e^{j\theta(h,K)} \right|^2 .
\tag{3.10}
$$

For any set of N complex numbers z_i, the following identity is valid:

$$
\left| \sum_{i=1}^{N} z_i \right|^2 = \sum_{i=1}^{N} |z_i|^2 + \sum_{i=1}^{N} \sum_{j=1, j\neq i}^{N} |z_i| \, |z_j| \, cos(\theta_i - \theta_j),
\tag{3.11}
$$

where θ_i is the argument of the complex number z_i in its exponential form $z_i = |z_i|\,e^{j\theta_i}$. Hence, Equation 3.10 can be written as follows:

$$P_{0_{M,K}} = \sum_{h \in N_{UE_K}\backslash\{M\}} P_R(h,K) + \sum_{p \in N_{UE_K}\backslash\{M\}} \sum_{q \in N_{UE_K}\backslash\{M,p\}} \sqrt{P_R(p,K)}\sqrt{P_R(q,K)}cos(\Delta\theta_{p,q}),$$
(3.12)

with $\Delta\theta_{p,q} = \theta(p,K) - \theta(q,K)$ being the phase difference between the two preambles using pattern (p, K) and (q, K), respectively.

The terms $P_R(h,K)$ make finding a closed form for $P_{0_{M,K}}$ complex. However, the satellite environment can ease the introduction of approximations. If the cell radius is much shorter than the satellite height, the slant range (i.e., the UE-satellite distance) is similar to the satellite height independently of the UE's position. For example, considering a certain satellite height h_{sat} and a cell radius $R_{cell} = 10\% \times h_{sat}$, the slant range has minimum value $d_{min} = h_{sat}$ (it is the case of a UE in the sub-satellite point, i.e. at the cell center) and maximum d_{max} (UE at the cell edge) as follows:

$$d_{max} \approx \sqrt{h_{sat}^2 + R_{cell}^2} = h_{sat}\sqrt{1 + \left(\frac{R_{cell}}{h_{sat}}\right)^2} \approx 0.5\% \times d_{min}.$$

This can be the case for a multi-beam coverage, where each satellite can cover a large area with several beams, effectively partitioning the coverage area into multiple small cells. Moreover, under this assumption the additional losses (e.g., atmospheric losses) are also similar. This means that the path loss applied to each preamble is approximately the same and, if the UEs use the same transmit power during NPRACH, the resulting received power is also similar. Naming this value P_R, the following equation can be obtained:

$$P_{0_{M,K}} = P_R\left(\sum_{h \in N_{UE_K}\backslash\{M\}} 1\right) + P_R \sum_{p \in N_{UE_K}\backslash\{M\}} \sum_{q \in N_{UE_K}\backslash\{M,p\}} cos(\Delta\theta_{p,q}) =$$

$$= P_R\left(|N_{UE_K} \backslash \{M\}| + \sum_{p \in N_{UE_K}\backslash\{M\}} \sum_{q \in N_{UE_K}\backslash\{M,p\}} cos(\Delta\theta_{p,q})\right) =$$
(3.13)

$$= P_R\left(|N_{UE_K}| - N_{UE_{M,K}} + \sum_{p \in N_{UE_K}\backslash\{M\}} \sum_{q \in N_{UE_K}\backslash\{M,p\}} cos(\Delta\theta_{p,q})\right),$$

where $|N_{UE_K}|$ represents the cardinality of the set N_{UE_K}.

Being $\Delta\theta_{p,q}$ the difference of two RVs, it is a RV itself; the same holds true when the cosine function is applied to it. Therefore, the mean value of $cos(\Delta\theta_{p,q})$ can be estimated

by averaging all of the possible values that the RV can assume. Hence, if $|N_{UE_K}|$ is large enough, the following approximation can be made:

$$\sum_{p \in N_{UE_K} \setminus \{M\}} \sum_{q \in N_{UE_K} \setminus \{M,p\}} \cos(\Delta\theta_{p,q}) \approx$$

$$\approx (|N_{UE_K}| - N_{UE_{M,K}})(|N_{UE_K}| - N_{UE_{M,K}} - 1) \times E[\cos(\Delta\theta_{p,q})] \tag{3.14}$$

In order to compute $E[\cos(\Delta\theta_{p,q})]$, the probability density function (PDF) of $\Delta\theta_{p,q}$ must be obtained first.

Introducing $\theta_1, \theta_2 \sim U[0, 2\pi]$, their joint PDF can be written as:

$$f_{\theta_1,\theta_2} = \begin{cases} \frac{1}{(2\pi)^2} & \text{if } \theta_1 \in [0, 2\pi] \text{ and } \theta_2 \in [0, 2\pi] \\ 0 & \text{otherwise} \end{cases}$$

Then, naming $\Delta\theta = \theta_1 - \theta_2$, it can be proven that its cumulative distribution function (CDF) is:

$$F_{\Delta\theta}(\xi) = \begin{cases} 0 & \text{if } \xi < -2\pi \\ \frac{1}{2}\frac{\xi^2}{4\pi^2} + \frac{\xi}{2\pi} + \frac{1}{2} & \text{if } \xi \in [-2\pi, 0] \\ -\frac{1}{2}\frac{\xi^2}{4\pi^2} + \frac{\xi}{2\pi} + \frac{1}{2} & \text{if } \xi \in [0, 2\pi] \\ 1 & \text{if } \xi > 2\pi \end{cases}$$

After computing the derivative of this formula with respect to ξ, the PDF of $\Delta\theta$ is obtained:

$$f_{\Delta\theta}(\xi) = \begin{cases} \frac{\xi}{4\pi^2} + \frac{1}{2\pi} & \text{if } \xi \in [-2\pi, 0] \\ -\frac{\xi}{4\pi^2} + \frac{1}{2\pi} & \text{if } \xi \in [0, 2\pi] \\ 0 & \text{otherwise} \end{cases}$$

However, recalling that the phase has periodicity 2π, the previous PDF can simply be written as:

$$f_{\Delta\theta}(\xi) = \frac{1}{2\pi} \text{ for } \xi \in [-\pi, \pi], \tag{3.15}$$

meaning that the difference between two phases uniformly distributed with support $[0, 2\pi]$ is still a uniformly distributed RV with the same support (shifted by π exploiting the phase's periodicity).

Finally, the PDF of the RV $Y = \cos(\Delta\theta)$ can be obtained through the following formula:

$$f_Y(y) = \frac{1}{\pi \sin(\cos^{-1} y)} \text{ for } y \in [-1, 1], \tag{3.16}$$

From the plot reported in Figure 3.4, it can be observed that the computed PDF tends
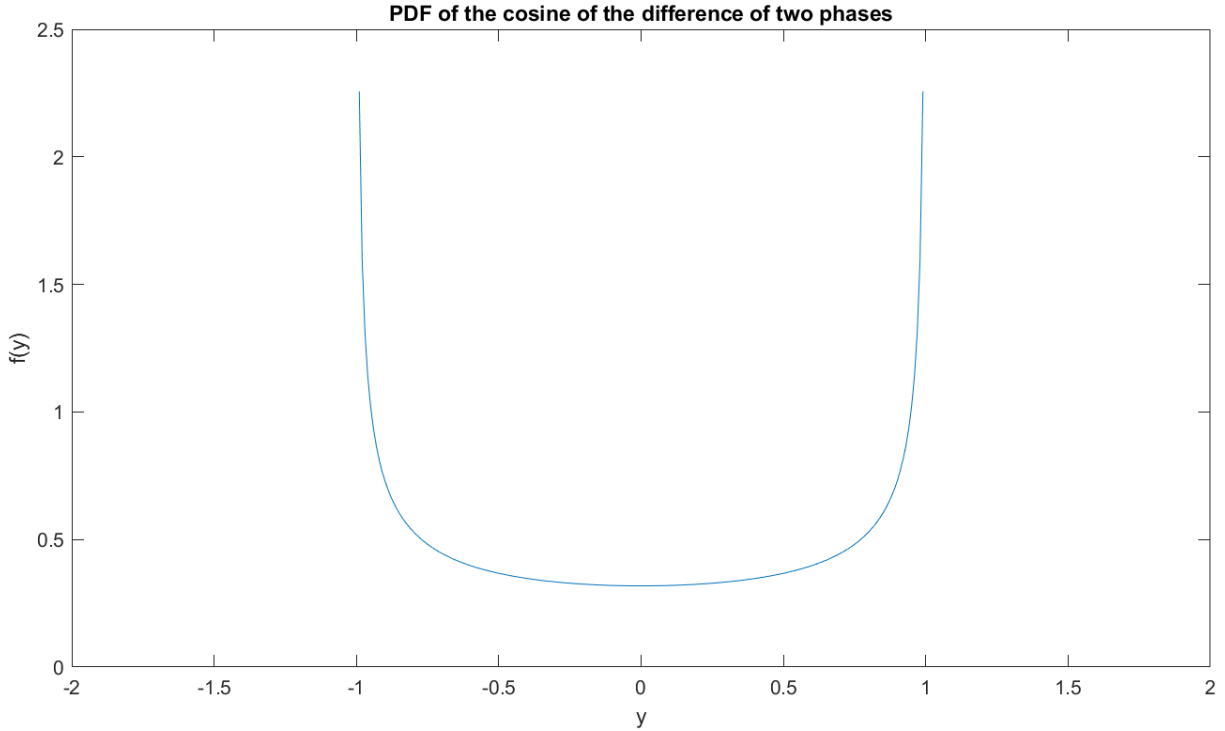
Figure 3.4: PDF of $Y = cos(\Delta\theta)$.

to infinite for $y = \pm 1$ and, most importantly, is symmetric in $y = 0$. From this last consideration, it is easy to state that $E[cos(\Delta\theta_{p,q})] = 0$. Hence, recalling the approximation introduced with Equation 3.14, Equation 3.13 reduces to:

$$P_{0_{M,K}} = P_R \left(|N_{UE_K}| - N_{UE_{M,K}} \right),$$

which can be written as follows using the definition in Equation 3.9:

$$P_{0_{M,K}} = P_R \left(\sum_{h=1, h \neq M}^{N_{SG}-1} N_{UE_{h,K}} \right), \text{ with } N_{UE_{M,K}} \in \{0, 1\}. \tag{3.17}$$

This means that the average received preamble power over the M-th SG of the K-th FH pattern directly depends on the number of users that are transmitting over that SG. From this consideration, the obtained formula can be generalized to more than one user per FH-ST preamble pattern:

$$P_{0_{M,K}} = P_R \left(\sum_{h=1, h \neq M}^{N_{SG}-1} N_{UE_{h,K}} \right), \text{ with } N_{UE_{M,K}} \in \mathbb{N}. \tag{3.18}$$

Finally, having determined the value of $P_{0_{M,K}}$, the distribution of the received power over the M-th SG of the K-th FH pattern is proven to be:

$$P_{M,K} \sim \chi_{10}^2(k, \gamma_0), \text{ with:} \tag{3.19}$$
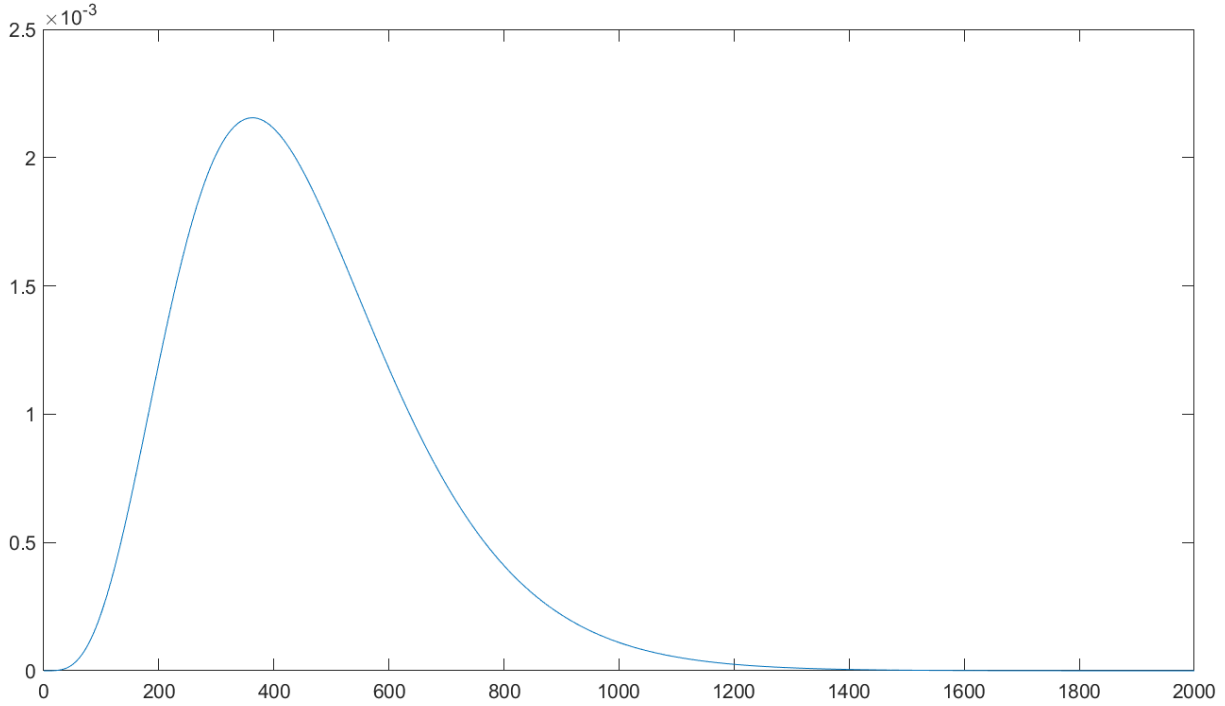
Figure 3.5: Example of PDF of $P_{M,K}$.

$$\text{Scale factor } k = \frac{10}{P_n} = \frac{10}{L \times N_0},$$

$$\text{Noncentrality parameter } \gamma_0 = 10\frac{P_{0_{M,K}}}{P_n} = 10\frac{P_R \left(\sum_{h=1,h\neq M}^{N_{SG}-1} N_{UE_{h,K}}\right)}{L \times N_0}.$$

An example of the PDF of $P_{M,K}$ is reported in Figure 3.5 assuming $P_R = N_0 = 0.875$ (estimated through simulation by considering a unitary preamble amplitude and $\frac{P_R}{N_0} = 1$) and $\sum_{h=1,h\neq M}^{N_{SG}-1} N_{UE_{h,K}} = 6$. A first issue of a NORA-based detector becomes evident, as $P_{M,K}$ can assume an extremely wide range of values. Moreover, it should be kept in mind that the received power over the M-th SG of the K-th FH pattern does not depend on the presence of a preamble (M, K) (as this preamble will silence the M-th SG, i.e., it will not contribute to the power received over the said SG). This means that the detection of the preamble (M, K) should be carried out by analyzing the power received over every SG of the K-th FH pattern *but the M-th.* As a consequence, a more complex detection algorithm should be implemented.

Recalling Equation 3.18, the received power over the M-th SG in absence of noise can be approximated as the sum of the power of each non-M-th preamble over the said SG. Hence, by introducing the ST pattern, the equation computed on the M-th SG differs

from the same equation written for the N-th SG by just a term in the sum (as the formula for SG M will have the power from the preamble (N, K) but not from the preamble (M, K), and vice versa). This system of equations can then be solved to determine which preamble is present, i.e., which preamble is associated to a non-negligible amount of received power. More specifically, a power threshold equal to the noise power was set to determine whether a preamble is present (corresponding power higher than threshold) or not.

It must be noted that, theoretically, the same algorithm could be applied to the standard NPRACH preamble. However, the ST pattern enables the addressability of the collided preambles. Moreover, this solution is expected to make the system of equations be solvable faster and with better precision.

However, depending on the SNR level, the additive noise can harshly affect the performance of the receiver. Even more importantly, the algorithm may not work as intended even in ideal noise-less conditions due to the approximations made: indeed, being the number of colliding preambles finite, Equation 3.14 may introduce errors in the detection process, which would ultimately result in performance degradation. In the next Chapter the proposed Energy Detector algorithm will be tested by means of simulation, presenting the estimated performances.

3.6 An alternative algorithm for NORA receivers

As already mentioned, the Energy Detector-based receiver may lead to unsatisfactory results. In search of improved performance, an alternative algorithm which does not rely on the approximation reported in Equation 3.14 was determined: a fitting-based algorithm. Recalling the mathematical representation of the sampled superimposed preambles at the receiver (Equation 3.5), each sample can be represented as a sum of a finite number of sinusoids, each randomly shifted independently of the other. Assuming that no more than one UE can choose the same preamble (i.e., each combination of FH and ST pattern can be chosen by one UE at most), this equation has 14 unknowns (without taking noise into account): fixing the FH pattern, for each of the 7 possible preambles the BS has no information on the corresponding received power (which, in case that no UE transmitted the preamble, is zero) and phase shift. However, these parameters remain the same for the entire duration of the RA opportunity. Hence, considering a large enough number of samples, the BS can estimate these parameters by fitting said equations (one per sample). To improve the algorithm performance, the received power per silenced preamble is assumed to be known at the BS: as stated in Section 3.4, the NTN scenario makes this

value approximately independent of the UE position inside the cell. Hence, each equation to be fitted can be written as:

$$r_{m,i}(n) = \sqrt{P_R} \sum_{h=1}^{N_{SG}-1} U(h) \times a_h(m) \times e^{j(2\pi \times f_K(m)\frac{n}{L} + \theta(h,K))} + w(n), \qquad (3.20)$$

with fitting parameters $U(h) \in \{0, 1\}$ (which indicates the presence or absence of the preamble (h, K)) and $\theta(h, K) \in [0, 2\pi[$ (which represents the phase shift introduced by the propagation). Considering N_{samp} samples per SG, the number of equations to be fitted is $N_{samp} \times N_{SG}$. As N_{samp} increases, better detection performance are expected while having the drawback of potentially increasing the estimation time. However, having $\theta(h, K)$ as fitting parameter means that this algorithm may also provide an estimation of the ToA, which is usually carried out by the BS after the preamble detection. Having both the detection and the ToA estimation carried out in a single step would save time and energy at the BS with respect to the standard algorithm. Moreover, being the CP identical to not only the first symbol, but also to every other symbol in the corresponding SG, the fitting-based algorithm would not need for it to be removed: indeed, the CP ensures that the remaining part of the SG is not overlapped with other preambles with a differential delay shorter than the CP length. However, as just stated, the algorithm does not need the entire SG but just N_{samp} samples out of it. This means that the samples could be taken from the middle of the SG, while the remaining part of it could be considered as an extended CP. To make this consideration clearer, assume that a 6-symbols SG is at the BS and that N_{samp} samples are taken from the 3rd and 4th symbols at BS time: being the symbols equal to each other, the received preamble can be delayed by at most twice the symbol time T_s (i.e., the BS is sampling the 1st and 2nd symbols instead of the 3rd and 4th) without introducing detection errors. This implies that the maximum differential delay that the NPRACH preamble can handle with the proposed algorithm is much larger than the one with the energy detector, i.e., the CP length (1 T_s). Being this value directly related to the cell radius (larger cells imply larger differential delays), larger beams can be considered, further increasing the number of UEs served at the same time.

It must be noted that the presence of noise can introduce errors in the fitting operation. However, the resulting performance may be particularly different from those of the energy detector. A more detailed analysis of the algorithm performance will be provided in the next Chapter, reporting the results of the simulated detection of multiple superimposed preambles (with ST pattern) through the proposed technique.

Chapter 4

NORA performance

To verify the goodness of the proposed algorithms, extensive simulations have been conducted in MathWorks' computing environment MATLAB. The Key Performance Indicators (KPIs) considered were the probability of Correct Detection (CD) P_{CD} and the probability of False Alarm (FA) P_{FA}: the first is the probability that a preamble is detected given that a UE effectively transmitted it; the second is the probability that the algorithm mistakes noise for a preamble, i.e., it detects a specific preamble while no UE transmitted it. The two KPIs were estimated by implementing a Monte Carlo simulation, i.e., by assessing the frequency at which the events happened over a large number of detections.

To do so, a random number of NPRACH preambles with equal FH pattern but distinct ST pattern were generated, each one characterized by a different phase shift to simulate the effects of the propagation. For this analysis, the number of repetitions N_{rep} was fixed to 2. The preambles were then superimposed one with the other and summed to an AWGN sample. The variance of the noise process, which was needed to generate said noise signal, was obtained by assessing the received power corresponding to a single preamble and fixing an SNR value; then, the noise power was computed by dividing the preamble power by the chosen SNR value in linear units:

$$P_N = \frac{E\left[|r_{m,i}(n)|^2\right]}{SNR}$$

This was done for a wide range of SNR values: in this way, the performance of the two algorithms could be evaluated for different levels of SNR.

It must be noted that usually the SNR is assessed at the receiver, meaning that the useful signal power is determined on the received signal (i.e., the superimposed preambles). However, given that the number of preambles generated (which directly affects the received signal power) was random, it was considered more significant to provide a

fixed signal power value: this is needed to ensure that, for the same SNR value, noise is generated from the same random process (i.e., with the same variance). Hence, the signal power was computed on a single preamble.

From this common point, the two detection algorithms presented in Sections 3.5 and 3.6 were applied to the simulated waveform. A high number of Monte Carlo iterations were performed to reach the desired accuracy. The two probabilities P_{CD} and P_{FA} were finally estimated for each SNR value in the considered range by dividing the number of detected preambles by the number of generated or absent preambles, respectively.

4.1 Energy Detector results

In order to provide satisfactory results, this simulation was run a total of 88 million times per SNR value, considering the range [-40, 0] in dB. Exploiting MATLAB's Parallel Computing Toolbox, the executions were distributed among 12 cores running in parallel to minimize the duration of the simulation. Overall, the entire set of simulations approximately took 11 days to be ran. Being the detection algorithm executed 41 times per run (once per SNR value), each batch of 12 parallel executions took 3.2 ms. Hence, the energy detector algorithm can run at the BS over the 48 FH patterns in parallel in just a few ms.

Figure 4.1 shows the P_{CD} graph as a function of the SNR. As reported in the plot, the CD probability increases with the SNR, starting from 5% at -40 dB. The curve finally reaches a plateau after SNR $\approx$ -10 dB, corresponding to a CD probability of about 70.4%. On the other hand, Figure 4.2 reports the behavior of P_{FA} as the SNR changes. Here, the curve has a similar trend with respect to the previous graph: at -40 dB the FA probability is around 2.2%, then it increases until it reaches 43.5% at SNR $\approx$ -5 dB. For larger SNR values, the FA probability is approximately constant. It must be noted that, despite the large number of executions, the curve shows spurious effects for SNR values lower than -20 dB. However, being the performance at higher SNR anything but satisfactory, it was concluded that further simulations to improve the estimated P_{FA} curve would not have provided any additional insights.

Overall, the performance can be said to be far from acceptable at any SNR level. Indeed, in order to obtain a better CD probability, P_{FA} must increase too. However, this does not mean that no energy detector-based algorithm can provide satisfactory results, as different versions of the algorithm could result in better performance. Nonetheless, it was determined that more efforts would have been better spent on the more promising
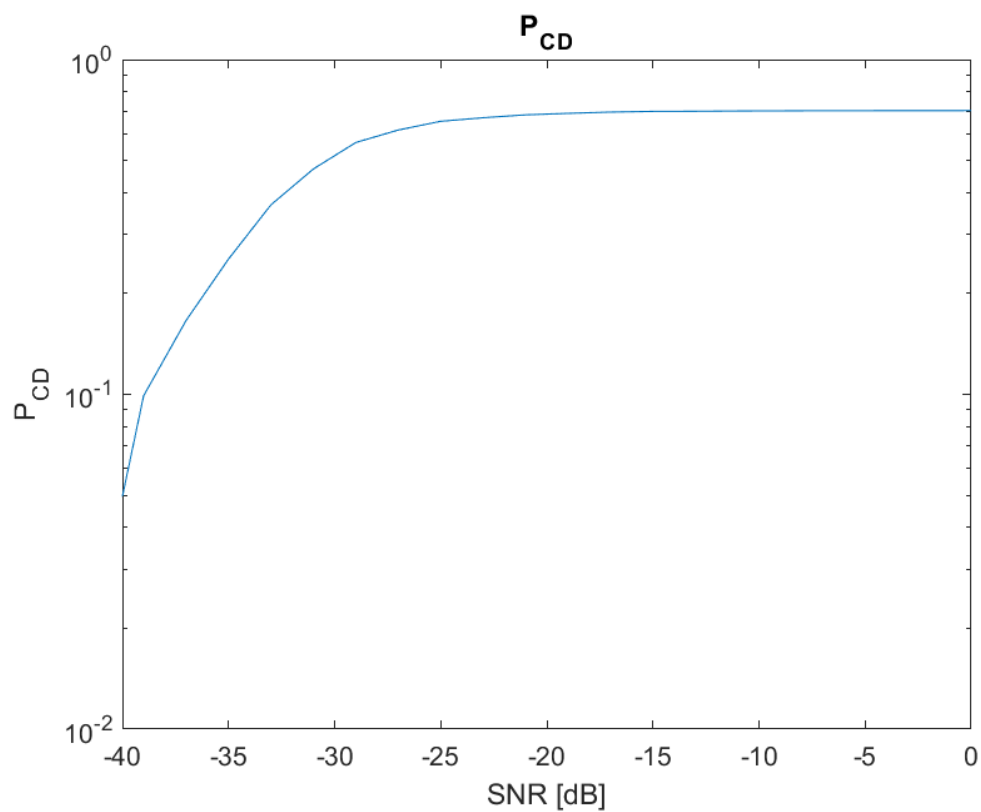
Figure 4.1: Probability of Correct Detection with the Energy Detector.
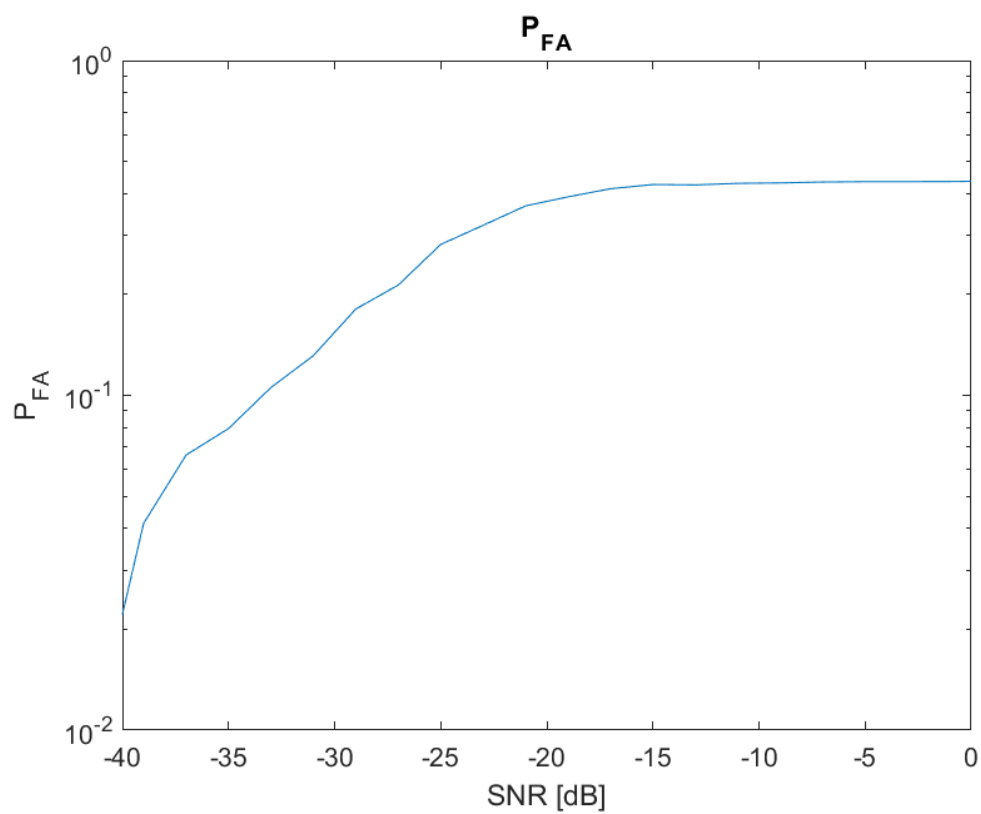

Figure 4.2: Probability of False Alarm with the Energy Detector.

fitting-based algorithm.

4.2   Fitting-based algorithm results

Finally, the fitting-based algorithm was tested in a separate set of simulations. In order to better determine the performance of the detector, one simulation was run for each possible number of superimposed preambles, i.e., N_p from 0 to 7. Each simulation consisted of 40'000 run per SNR value. The SNR range [-20, 20] dB was chosen, which was considered to be a realistic setup for satellite communications. For these simulations, a server-grade CPU was needed: it was estimated that the 12 cores CPU mentioned in Section 4.1 would have taken approximately 4 days to execute only one of the eight simulations, without taking into account the available RAM and other hardware limitations. Hence, a 56 cores CPU was utilized, dropping the execution time to around 32 hours per simulation. With this setup, each set of 56 detections took approximately 4 seconds.

The 7 P_{CD} graphs are reported in Figures [4.3 - 4.9], while Figure 4.10 reports the said graphs overlapped. It must be noted that the curve is not reported for the simulation with 0 preambles: indeed, in this case no preambles are generated, hence no preambles can be correctly detected. However, from this observation it can be concluded that, as the number of superimposed preambles increases, the sample size for the estimation of the P_{CD} curve increases too: for example, in the simulation with $N_p = 6$ a maximum of 6 CDs can be made. For this reason, the simulations with $N_p = 1$ and $N_p = 2$ were performed with double the number of runs (i.e., 80'000): if this was not the case, the resulting P_{CD} graphs would have ended up being too inaccurate to report meaningful insights on the algorithm's performance.

At a first glance, the fitting-based algorithm promises much better CD capabilities than the ones offered by the energy detector: all of the curves start at around 82% for SNR = -20 dB. As a comparison, the energy detector was able to reach a maximum CD probability of 70.4%, requiring an even higher SNR to reach this value. At SNR = 20 dB the curves reach at least 88% (corresponding to $N_p = 6$). The performance of the fitting-based detector seems to slightly decrease as N_p increases: albeit within margin of error, this behavior is to be expected, as the presence of multiple preambles can make the detection more difficult. In the best case (i.e., $N_p = 1$), P_{CD} at SNR = 20 dB is just over 91%. However, Figures [4.3 - 4.9] show an unexpected behavior of the P_{CD} curve: the plots report lower P_{CD} performances for SNR values around 0 dB for any N_p value. Even though the performance drop is not critical, as P_{CD} remains mostly over 80%, the fact that it is
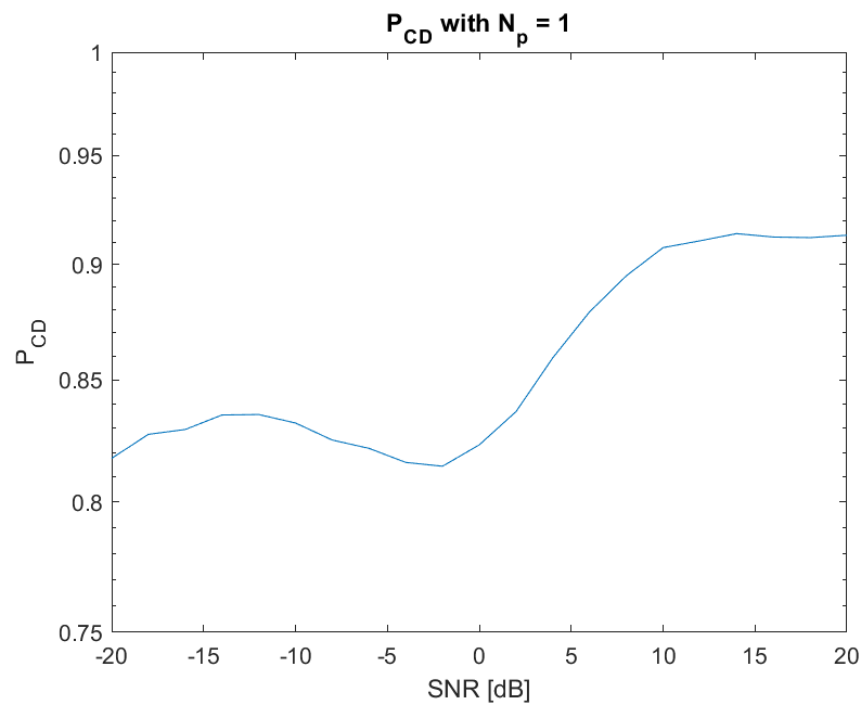
Figure 4.3: Probability of Correct Detection with the fitting algorithm ($N_p = 1$).
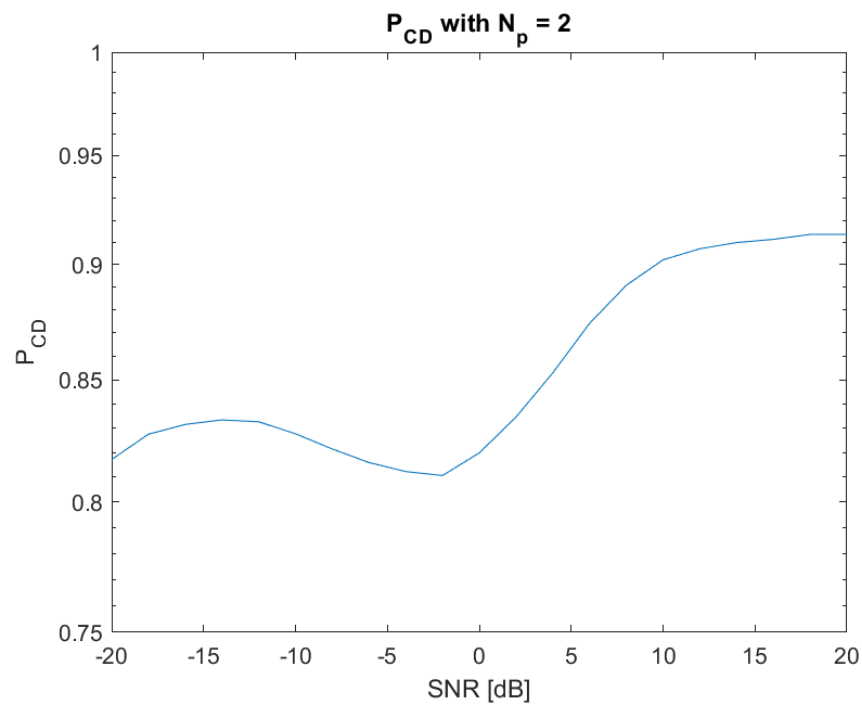

Figure 4.4: Probability of Correct Detection with the fitting algorithm ($N_p = 2$).
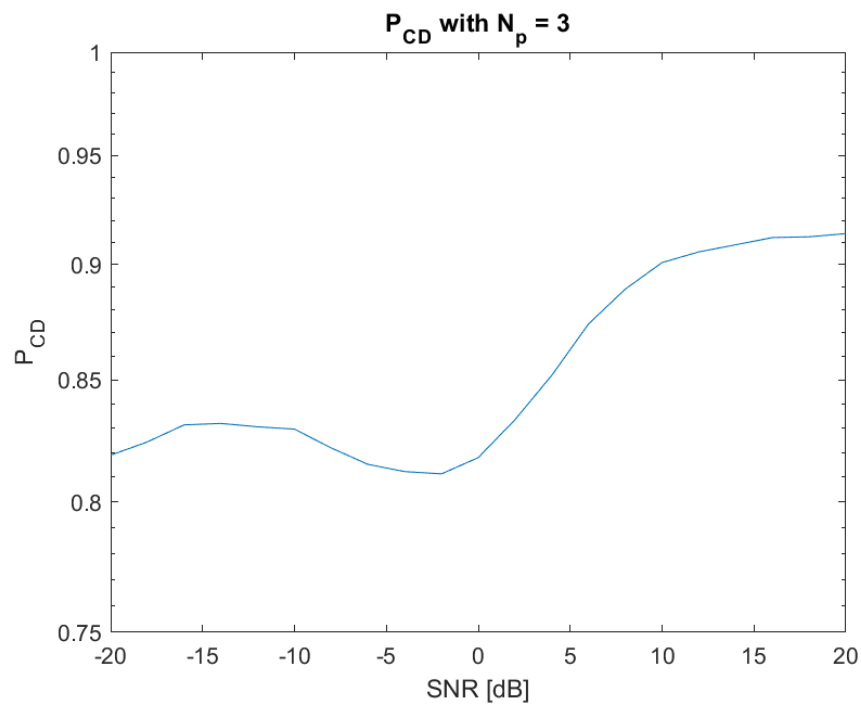
Figure 4.5: Probability of Correct Detection with the fitting algorithm ($N_p = 3$).
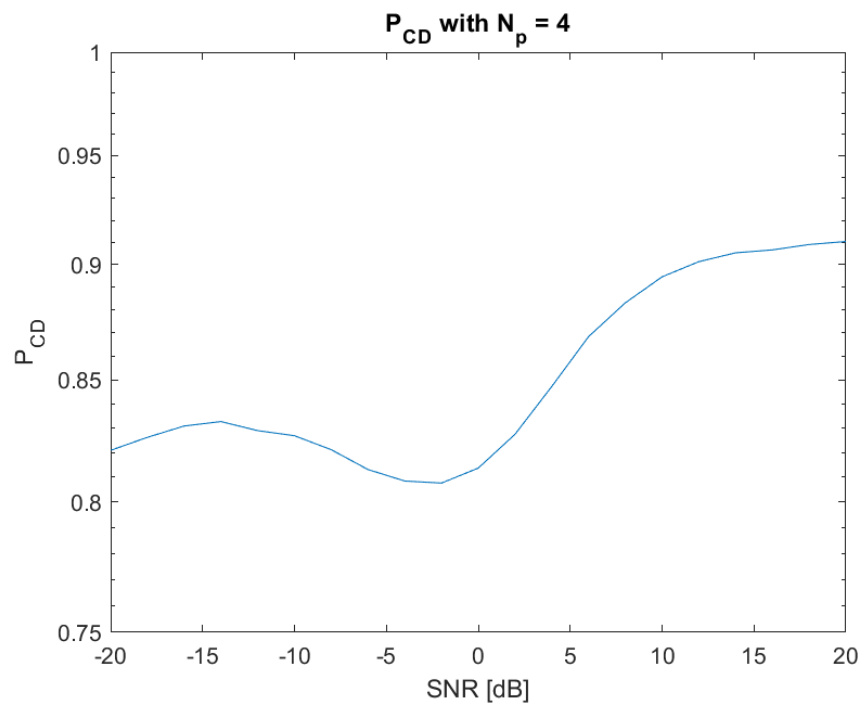

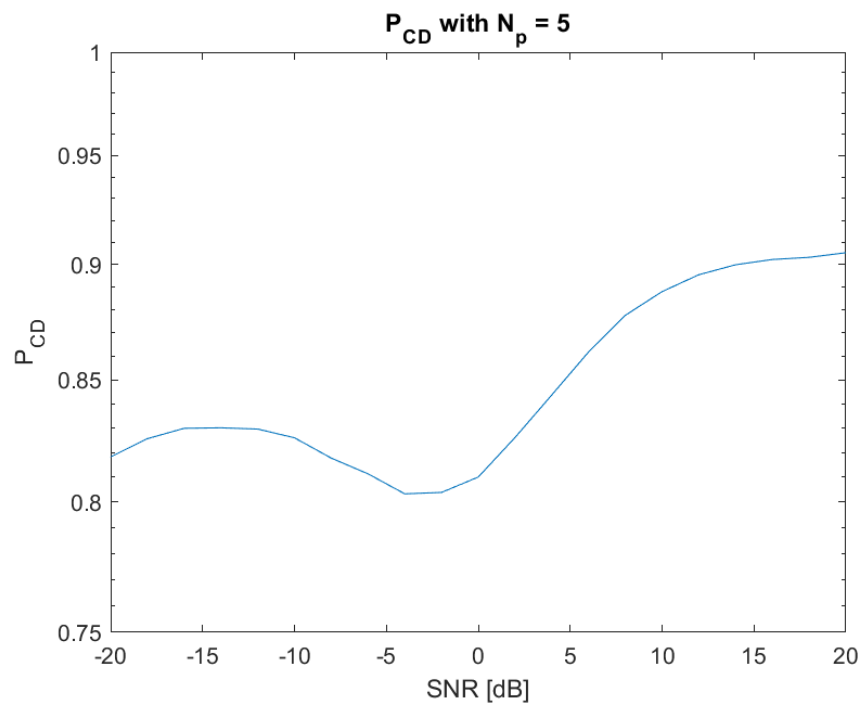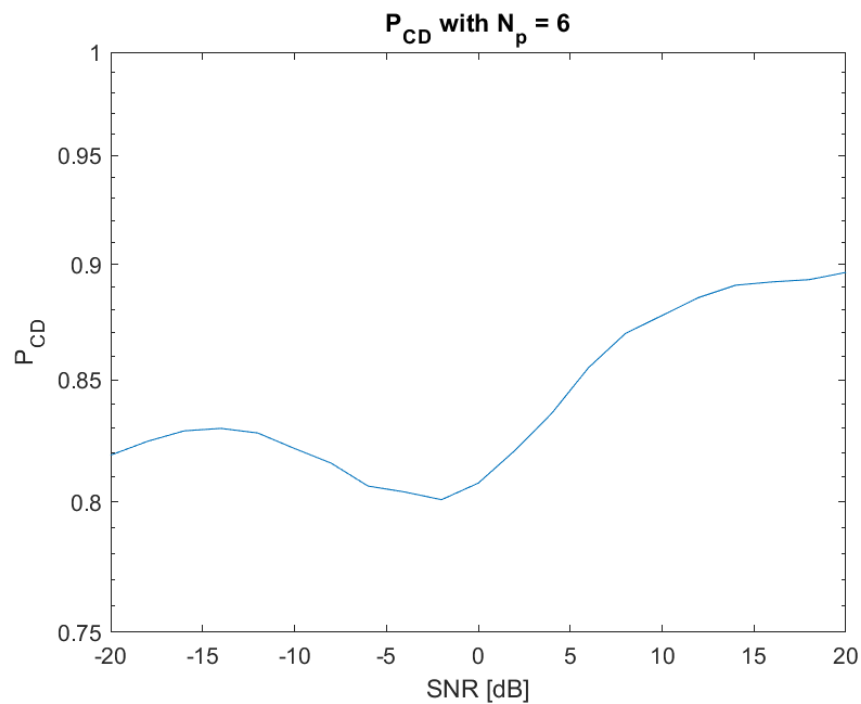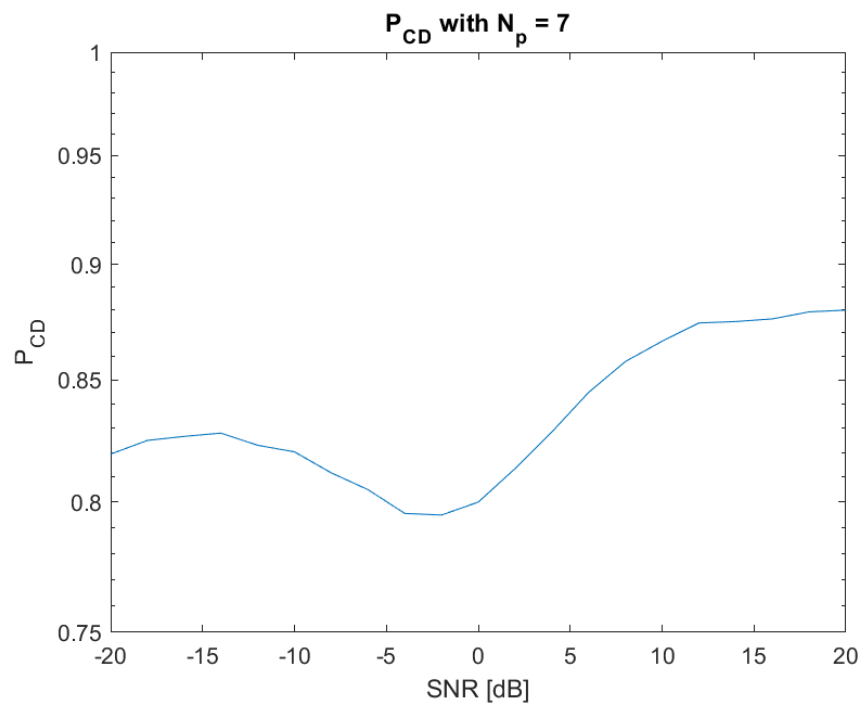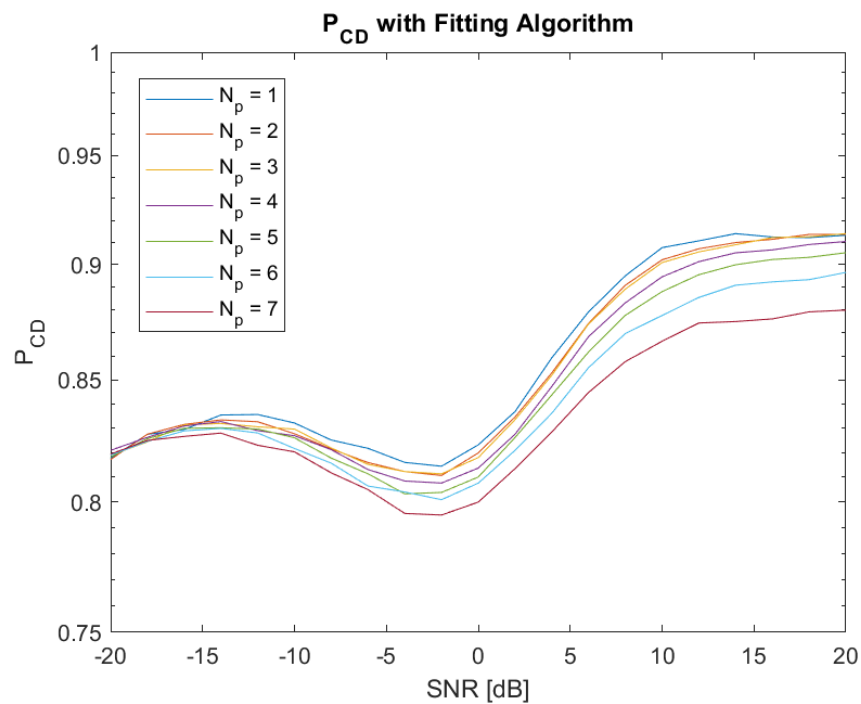Figure 4.6: Probability of Correct Detection with the fitting algorithm ($N_p = 4$).

Figure 4.7: Probability of Correct Detection with the fitting algorithm ($N_p = 5$).

Figure 4.8: Probability of Correct Detection with the fitting algorithm ($N_p = 6$).

Figure 4.9: Probability of Correct Detection with the fitting algorithm ($N_p = 7$).

Figure 4.10: Probability of Correct Detection with the fitting algorithm.

consistently reported in each simulation suggests that this behavior does not happen by chance. To determine the source of these missed detections, it is necessary to recall the working principle of the algorithm presented in Section 3.6: this algorithm tries to fit a set of equations, each of them being a sum of sinusoids at a constant frequency f_c. It is important to remember that the received power of each individual preamble (hence, their amplitude) is assumed to be equal to each other. Recalling that the SNR was computed as the ratio between the received power corresponding to a single preamble and the noise power, it is immediate to state that for SNR values close to 0 dB (i.e., 1 in linear units) the noise power is similar to the power of a single preamble. Being noise white, it can then be said that the amplitude of its frequency components close to f_c is similar to the amplitude of a received preamble. Hence, the reason why the CD probability of this algorithm is slightly lower around SNR = 0dB is probably because in this range noise can be particularly mistaken for a preamble, and vice versa.

The same considerations reported for P_{CD} can be made for P_{FA}, with the corresponding graphs presented in Figures [4.11 - 4.17] and overlapped in Figure 4.18. Indeed, here the simulation with $N_p = 7$ does not provide any performance indication on the FA probability, as all preambles are present and no FA can happen. As N_p decreases, the sample size of the simulation goes up, as the number of absent preambles to be wrongly detected increases. However, the same drop in performance reported in the P_{CD} graphs (here to be expected as a spike in FAs) is not present: this is most probably due to the fact that, in this range, the FA probability is already particularly high and the added effect of the noise can only contribute to prolong this bad performance region to higher SNR values. Moreover, differently from the P_{FA} plots presented in Section 4.1, the FA probability has now the opposite behavior with respect to the CD probability: the curve starts at SNR = -20 dB with a P_{FA} in the range [72%, 41%] and drops to [3.6%, 0.24%] at SNR = 20 dB. This means that, ideally, the NTN should be designed to work at the highest achievable SNR level to perform the best: given the plateauing behavior of both the curves for any N_p value, an SNR higher than 15 dB should be desired. Finally, differently from the P_{CD} plots reported in this Section, the FA performance increases with decreasing N_p only after SNR = 10 dB; before this threshold, the behavior is twisted, and the better performing curves are the ones corresponding to higher N_p.

As for the energy detector, this algorithm is far from perfect: as already reported, the simulation exploited the parallel computational capabilities of a 56 cores CPU, meaning that 56 detections were performed in parallel in under 4 seconds. Clearly, the excessively large duration of the detection is the main drawback of this algorithm. However, the
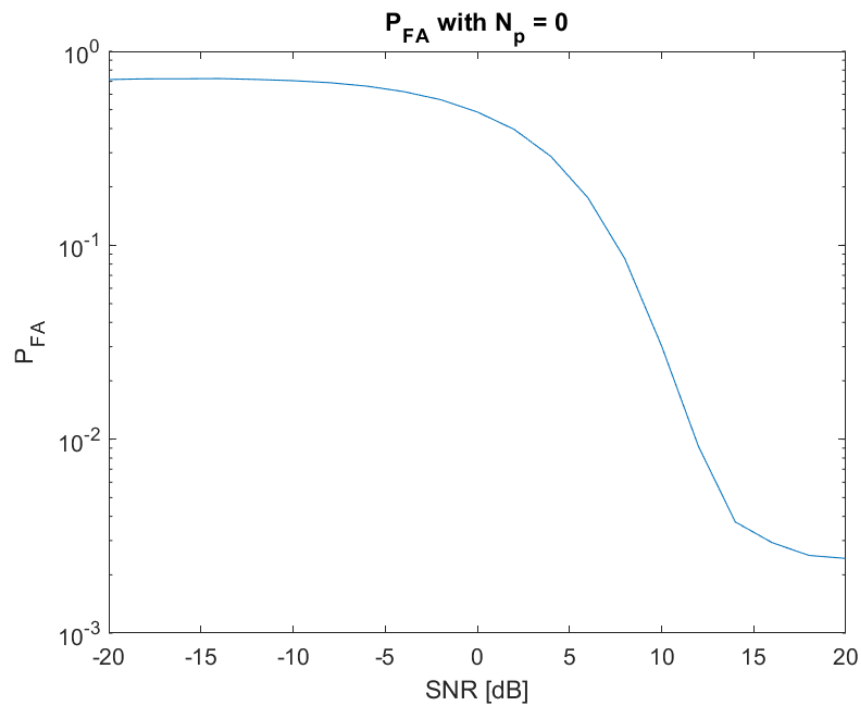
Figure 4.11: Probability of False Alarm with the fitting algorithm ($N_p = 0$).
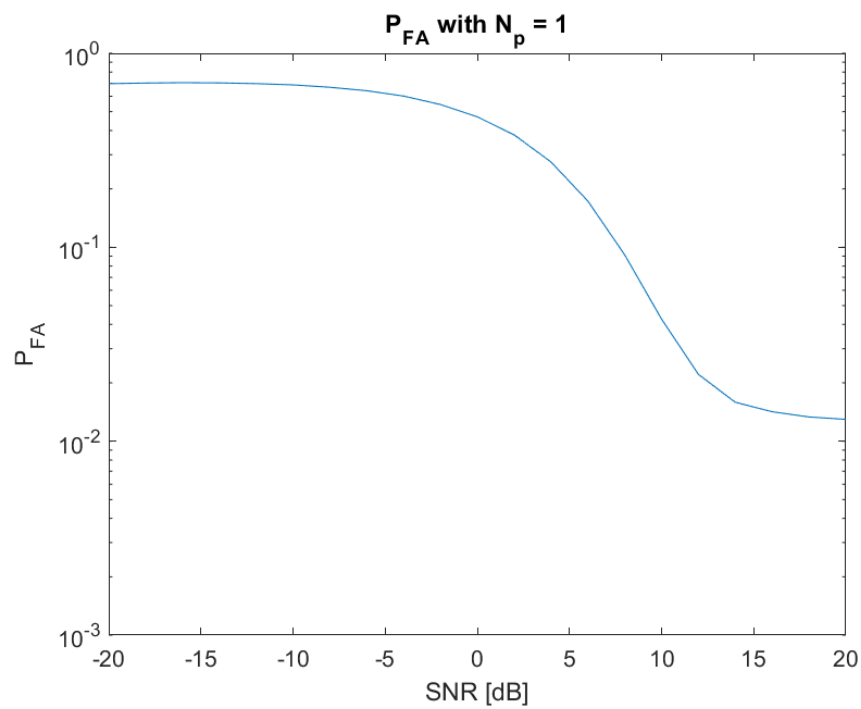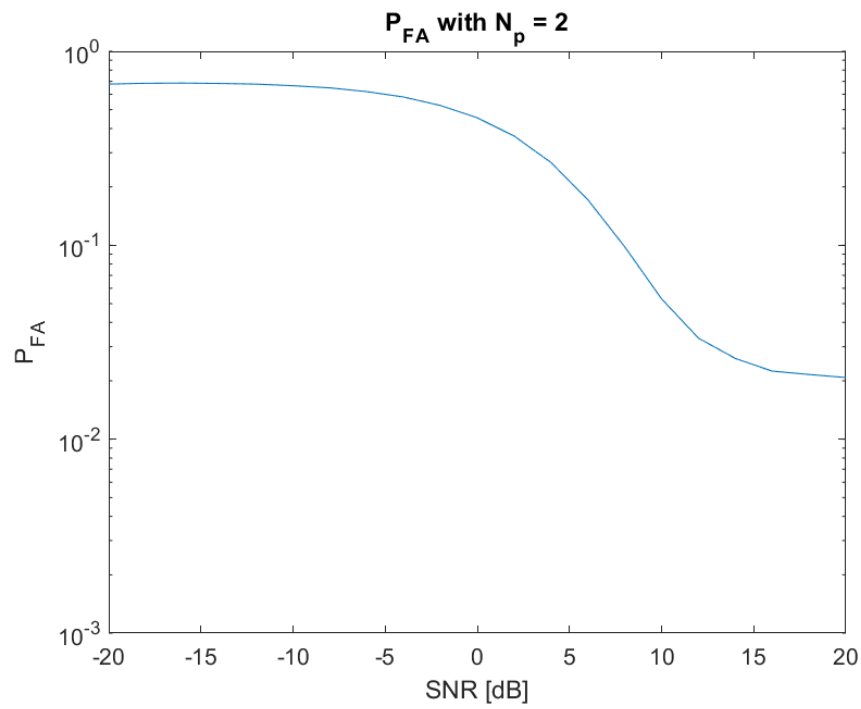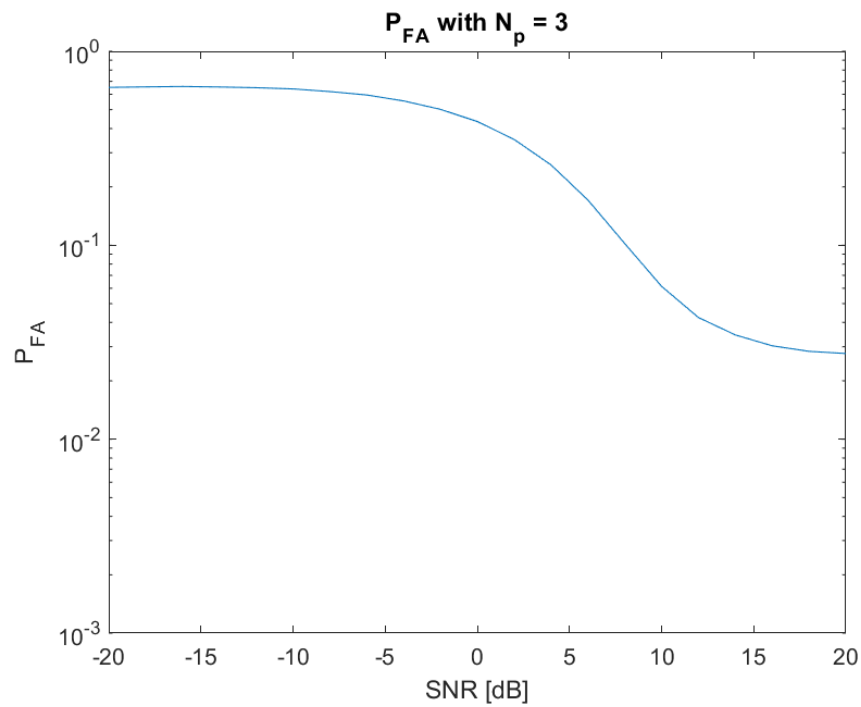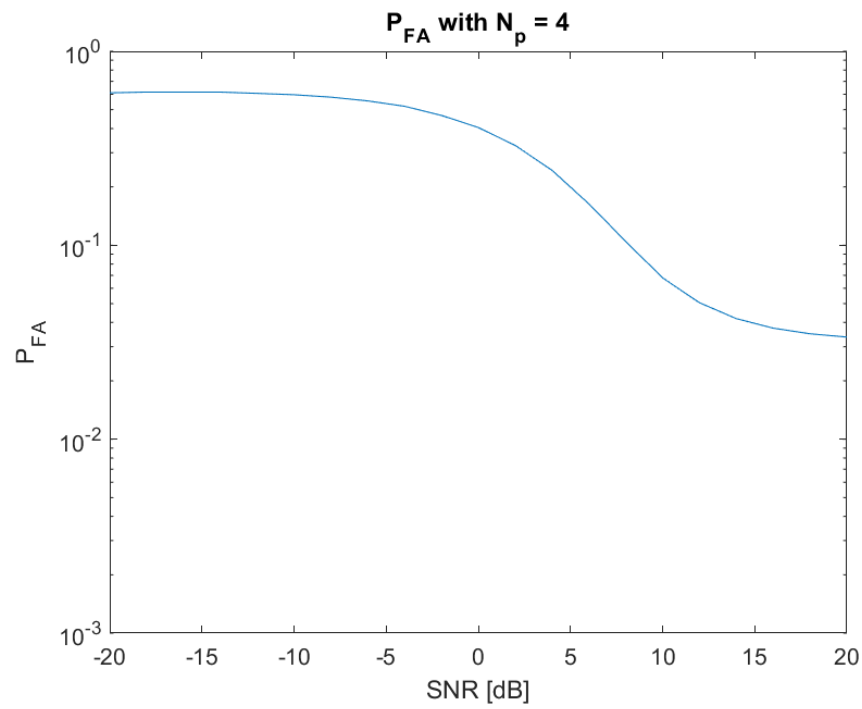

Figure 4.12: Probability of False Alarm with the fitting algorithm ($N_p = 1$).

Figure 4.13: Probability of False Alarm with the fitting algorithm ($N_p = 2$).

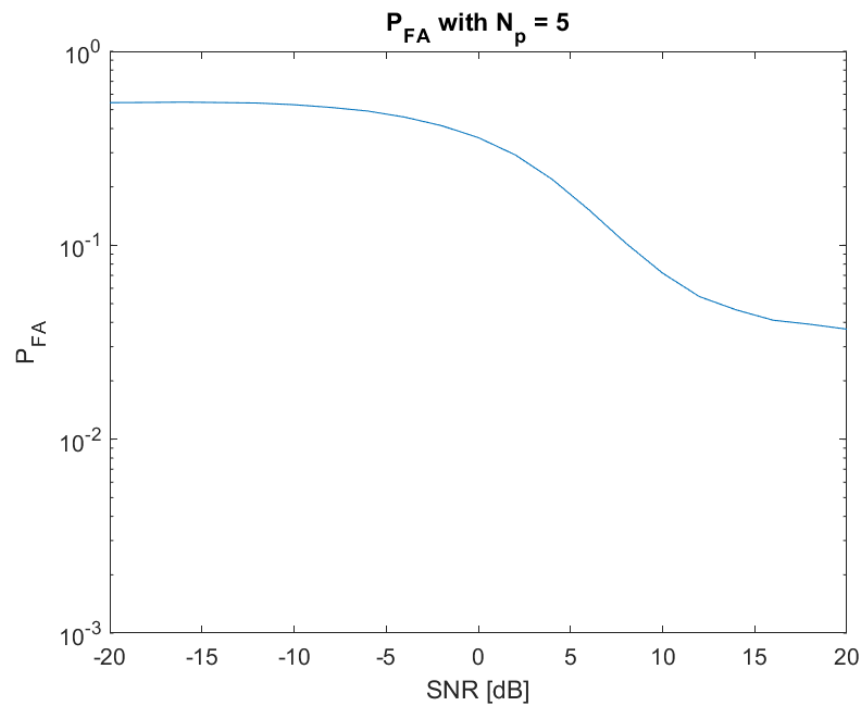

Figure 4.14: Probability of False Alarm with the fitting algorithm ($N_p = 3$).

Figure 4.15: Probability of False Alarm with the fitting algorithm ($N_p = 4$).

Figure 4.16: Probability of False Alarm with the fitting algorithm ($N_p = 5$).
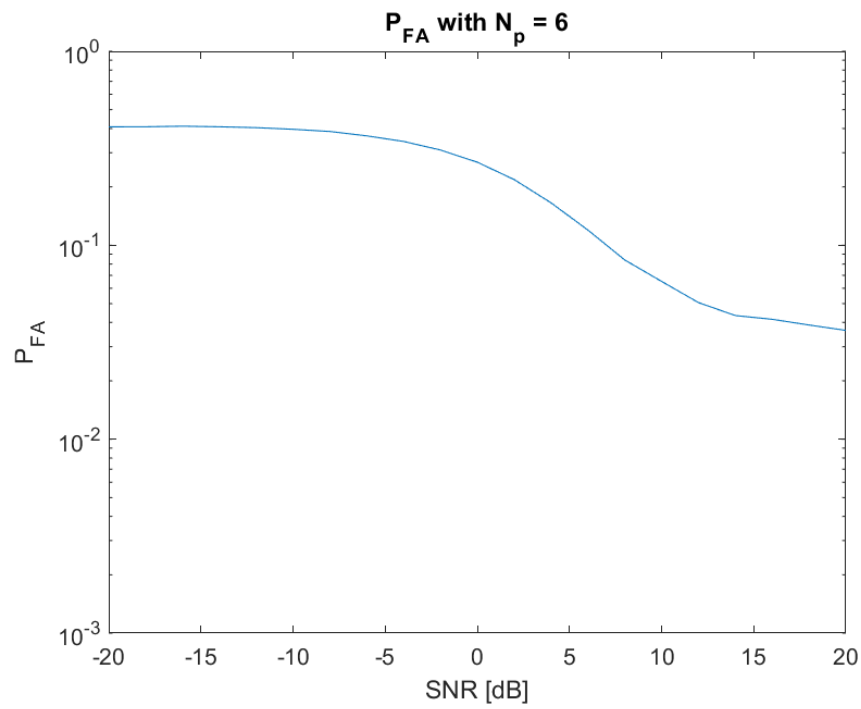
Figure 4.17: Probability of False Alarm with the fitting algorithm ($N_p = 6$).
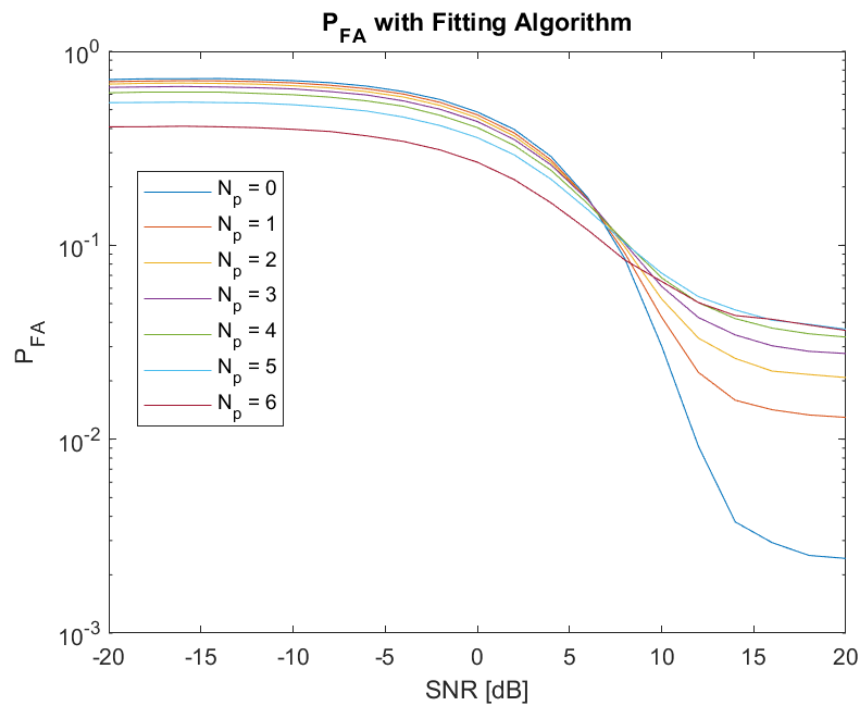

Figure 4.18: Probability of False Alarm with the fitting algorithm.

real implementation of the algorithm at the BS could be further optimized by choosing a different solver, being MATLAB's *fmincon* the one used in this work. Moreover, the detection algorithm should run on each of the 48 distinct FH patterns in parallel, meaning that 48 cores in the BS CPU would have to be reserved for the NPRACH detection during each RA opportunity. Depending on the BS position (i.e., on the satellite or on ground), this requirement could affect the overall system's complexity. Nonetheless, the performance of the proposed algorithm are overall promising, bringing NORA one step closer to being reality.

Conclusions

This thesis presented a Non-Orthogonal approach to the Random Access procedure, taking as a case study a Narrowband-IoT Non-Terrestrial Network. After having introduced the said wireless access standard, the Random Access procedure was analyzed in depth to highlight its shortcomings with respect to the envisioned Massive IoT era. A slight modification of the NPRACH preamble was then introduced, called Silenced Tone pattern, in order to enable NORA. With the new waveform in mind, the statistical characterization of the received power of superimposed NORA NPRACH preambles was carried out in order to introduce an energy detector-based receiver. An alternative algorithm, based on equation fitting, was also presented. Finally, the performance of the two algorithms was estimated by means of simulation, reporting the Correct Detection and False Alarm probabilities as a function of the SNR. Although the fitting-based receiver was determined to be the better performing among the two, it is characterized by a long execution time. However, this may change significantly depending on the implementation, i.e., on the solver used, the hardware of the BS, et cetera.

As mentioned in Section 3.6, the fitting-based algorithm is capable of performing a ToA estimation, too. Even though it was out of the scope of this work, further analysis could be carried out in the future to determine the accuracy of these estimates. If the estimation performance turned out to be satisfactory, the ToA could be estimated by the BS without any additional time or energy needed, as it would be implicitly carried out during the NPRACH preamble detection. Other future steps could foresee the introduction of a neural network-based or correlation-based receiver in an effort to reduce the execution time of the detection. Furthermore, an analysis of the impact of typical NTNs impairments (i.e., large delays and Doppler shifts) on the detection performance could be carried out. For what regards the energy detector-based algorithm, different variants could be experimented, e.g., simply changing the threshold to a noise-independent value or taking into account the presence of noise in the system as a further unknown. Finally, system-level simulations could be done in order to determine further KPIs, such as the system throughput, the average number of RA attempts per UE and the average number of UEs that are not able to connect during the entire satellite's visibility window.

Bibliography

[1] Liberg O., Sundberg M., Wang Y.-P. E., Bergman J., and Sachs J., *Cellular Internet of Things. Technologies, Standards, and Performance*, Elsevier Ltd., 2nd edition, 2020.

[2] Houcine Chougrani, Steven Kisseleff, Wallace A. Martins, and Symeon Chatzinotas, "Nb-iot random access for non-terrestrial networks: Preamble detection and uplink synchronization," *IEEE Internet of Things Journal*, pp. 1–1, 2021.

[3] 3GPP, *TR 45.820, Cellular system support for ultra-low complexity and low throughput Internet of Things (CIoT)*, 2015.

[4] 3GPP, *TR21.915 Release description; Release 15*, 2019.

[5] Ashton K., "That 'internet of things' thing," *RFID Journal*, 2009.

[6] Cisco, "Cisco annual internet report (2018–2023) white paper," 2020, `https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html`.

[7] Gartner, "Forecast: It services for iot, worldwide, 2019-2025," 2021, `https://www.gartner.com/en/documents/4004741#:~:text=Summary,occur%20in%20the%20run%20phase`.

[8] 3GPP, *TR 22.868, Study on Facilitating Machine to Machine Communication in 3GPP Systems*, 2007.

[9] ECC, *THE EUROPEAN TABLE OF FREQUENCY ALLOCATIONS AND APPLICATIONS IN THE FREQUENCY RANGE 8.3 kHz to 3000 GHz (ECA TABLE)*, 2019.

[10] FCC, *FCC ONLINE TABLE OF FREQUENCY ALLOCATIONS*, 2020.

[11] LoRA Alliance Technical Marketing Workgroup 1.0, *LoRaWAN, What Is It? A Technical Overview of LoRa and LoRaWAN*, 2015.

[12] ETSI Industry Specification Group on Low Throughput Networks, *GS LTN 003 V1.1.1 Low Throughput Networks (LTN); Protocols and Interfaces*, 2014.

[13] Ingenu, *RPMA Technology for Internet of Things*.

[14] Harwahyu R., Cheng R.-G., Tsai W.-J., Hwang J.-K., and Bianchi G., "Repetitions versus retransmissions: Tradeoff in configuring nb-iot random access channels," *IEEE INTERNET OF THINGS JOURNAL*, 2019.

[15] Parkvall S. Dahlman E. and Sköld J., *4G, LTE-Advanced Pro and The Road to 5G*, Elsevier Ltd., 3rd edition, 2016.

[16] Seongchul Cho, Hyungjin Kim, and Gweondo Jo, "Determination of optimum threshold values for nprach preamble detection in nb-iot systems," in *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, 2018, pp. 616–618.

[17] Lin X., Adhikary A., and Wang Y.-P. E., "Random access preamble design and detection for 3gpp narrowband iot systems," *IEEE WIRELESS COMMUNICATIONS LETTERS*, 2016.

[18] Hwang J.-K., Li C.-F., and Ma C., "Efficient detection and synchronization of superimposed nb-iot nprach preambles," *IEEE INTERNET OF THINGS JOURNAL*, 2019.

[19] Chakrapani A., "Nb-iot uplink receiver design and performance study," *IEEE INTERNET OF THINGS JOURNAL*, 2019.

[20] Chenchen Zhang, Nan Zhang, Wei Cao, Kaibo Tian, and Zhen Yang, "A design of nb-iot random access preamble receiver for large frequency offset," in *2021 ITU Kaleidoscope: Connecting Physical and Virtual Worlds (ITU K)*, 2021, pp. 1–6.

[21] Carla Amatetti, Matteo Conti, Alessandro Guidotti, and Alessandro Vanelli-Coralli, "Preamble detection in nb-iot via satellite: a wavelet based approach," in *2021 IEEE Global Communications Conference (GLOBECOM)*, 2021, pp. 1–6.

[22] Seyedi Y. and Safavi S. M., "On the analysis of random coverage time in mobile leo satellite communications," *IEEE Communications Letters*, 2012.

[23] Seokjae Moon, Hyun-Suk Lee, and Jang-Won Lee, "Sara: Sparse code multiple access-applied random access for iot devices," *IEEE Internet of Things Journal*, vol. 5, no. 4, pp. 3160–3174, 2018.

[24] Han Seung Jang, Hoon Lee, and Tony Q. S. Quek, "Deep learning-based power control for non-orthogonal random access," *IEEE Communications Letters*, vol. 23, no. 11, pp. 2004–2007, 2019.

[25] Yeduri Sreenivasa Reddy, Ankit Dubey, Abhinav Kumar, and Trilochan Panigrahi, "A successive interference cancellation based random access channel mechanism for machine-to-machine communications in cellular internet-of-things," *IEEE Access*, vol. 9, pp. 8367–8380, 2021.

[26] Yi Hou, Huiyue Yi, Wuxiong Zhang, and Hui Xu, "Nora based group paging overload control method for massive mtc accesses under 5g networks," in *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)*, 2021, pp. 692–697.

[27] D. Wang, Y. Qu, and Y. et al. Fu, "A non-orthogonal random access scheme based on nb-iot," in *Wireless Personal Communications*, 2020, vol. 111, p. 2625–2639.