Article

Analysis of Power Allocation for NOMA-Based D2D Communications Using GADIA

Husam Rajab [1,*], Fatma Benkhelifa [2] and Tibor Cinkler [1]

1   Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, 1111 Budapest, Hungary; Cinkler@tmit.bme.hu
2   Department of Computing, Faculty of Engineering, Imperial College London, London SW7 2BX, UK; f.benkhelifa@imperial.ac.uk
*   Correspondence: Husamrajab@tmit.bme.hu

Abstract: The new era of IoT brings the necessity of smart synergy for diverse communication and computation entities. The two extremes are, on the one hand, the 5G Ultra-Reliable Low-Latency Communications (URLLC) required for Industrial IoT (IIoT) and Vehicle Communications (V2V, V2I, V2X). While on the other hand, the Ultra-Low Power, Wide-Range, Low Bit-rate Communications, such as Sigfox, LoRa/LoRaWAN, NB-IoT, Cat-M1, etc.; used for smart metering, smart logistics, monitoring, alarms, tracking applications. This extreme variety and diversity must work in synergy, all inter-operating/inter-working with the Internet. The communication solutions must mutually cooperate, but there must be a synergy in a broader sense that includes the various communication solutions and all the processing and storage capabilities from the edge cloud to the deep-cloud. In this paper, we consider a non-orthogonal multiple access (NOMA)-based device to device (D2D) communication system coexisting with a cellular network and utilize Greedy Asynchronous Distributed Interference Avoidance Algorithm (GADIA) for dynamic frequency allocation strategy. We analyze a max–min fairness optimization problem with energy budget constraints to provide a reasonable boundary rate for the downlink to all devices and cellular users in the network for a given total transmit power. A comprehensive simulation and numerical evaluation is performed. Further, we compare the performance of maximum achievable rate and energy efficiency (EE) at a given spectral efficiency (SE) while employing NOMA and orthogonal frequency-division multiple access (OFDMA).

Keywords: IoT; 5G; eMBB; mMTC; URLLC; GADIA; NOMA; OFDMA

1. Introduction

Recently, a rising consensus in the fifth generation (5G) of wireless network technology has been developed to sustain the massive number of users. It will support a broad range of vertical industries demands by linking everyone and everything, particularly smart devices, machines, and sensors. The 5G network will drive innovation and has the prospect to change the way we live, work and play. According to the International Telecommunication Union (ITU), 5G wireless systems are classified to support three generic services as enhanced Mobile Broadband (eMBB), massive Machine-Type Communications (mMTC), and Ultra-Reliable and Low-Latency Communications (URLLC) [1,2]. These use cases empower various services provided by the 5G network: (a) the eMBB provides a stable connections service of mobile broadband with significantly high peak data rates that assist in providing consistent user experience; (b) the mMTC furnishes a massive number of Internet of Things (IoT) devices, which are transmitting small data payloads periodically; (c) the URLLC services will open up additional diverse capabilities of the network by expanding its limits for high reliability and low-latency from a restricted set of terminals. The miscellaneous deployment of the 5G network is due to the increase in various new characteristics, i.e., scalable numerology, along with compatibility, elastic

spectrum, and ultra-lean design. The current technologies of 5G structures form the basis for numerous new technologies beyond 5G/6G network, as Figure 1 illustrates. There are two approaches to sustain the coexistence of various services, i.e., orthogonal slicing and non-orthogonal slicing. In the orthogonal slicing approach, different resources are assigned to each user to satisfy their requirements. The allocated resources are orthogonal in frequency-time, i.e., no interference [3]. The significant drawback here is an inefficient use of the available resources. In the non-orthogonal slicing approach, the same resources are divided for all services simultaneously. Reusing similar resources enables improvement in spectral efficiency regardless of possible interference between these services [4,5].

The main purpose of eMBB service is to increase the data rate while maintaining reasonable reliability. Furthermore, an mMTC device is active periodically and utilizes a fixed low transmission rate in the uplink. Furthermore, mMTC has been developed as part of 3GPP Release 13/14 low power wide area (LPWA) technologies which includes Narrowband Internet of Things (NB-IoT). The enormous number of conceivable active mMTC devices results in difficulties to assign primary resources to particular mMTC devices. Instead, it is essential to provide resources that can be communicated via random access. For example, grant-free random access with an excessive number of non-orthogonal pilot sequences is introduced as a promising technique for supporting mMTC where all active devices (with unique identity) synchronously transmit their pilot sequences and the goal of an access point is to detect the activities of devices [6].

The size of mMTC devices' active division is a random variable whose average value estimates the mMTC traffic arrival rate. The primary purpose of designing the mMTC is to magnify the arrival rate that can be sustained in a shared radio resource. The targeted packet error probability (PER) of an individual mMTC transmission is typically low, e.g., in the order of 10^{-1}. Although the URLLC transmissions are periodic, the set of potential URLLC transmitters is much smaller than those of mMTC. For sustaining the URLLC communications, it requires a combination of scheduling algorithms to utilize the available resources for reliable data transmission efficiency. Moreover, to prevent too many resources from being idle due to the periodic traffic, identifying the transmission of URLLC on time because the low latency will require comprehensive coverage, which is highly unlikely in early 5G deployments. Enormous diversity is essential to high-reliable communication [7], and it can be accomplished only by utilizing multiple frequency or spatial resources. The transmission rate of a URLLC is relatively insufficient, and the main requirement is a high-reliability level (with a PER generally lower than 10^{-5}). That is undoubtedly a fundamental purpose that is considered by both academia and industry. Notably, wireless local area networks have enhanced increasingly popular over recent years, demanding massive connections between numerous devices and a base station and between devices termed as device-to-device communication (D2D) [8]. Due to the spectrum scarcity for cellular users, it requires more local facilities. Integrating D2D links within the existing cellular infrastructure can improve the spectrum energy efficiency, increase the network coverage, and enhance communication reliability [9]. Several techniques have been merged into a D2D scheme to provide more reliable communication in ultimate wireless networks, e.g., massive multiple-input multiple-output (MIMO) [10,11], nonorthogonal multiple access (NOMA) [12], and mmWave communications [13]. In particular, NOMA has recently gained significant attention as a promising multiple access applicant for superposition coding, and successive interference cancellation [14]. It is determined as NOMA in the power domain, and the network can assist multiple users simultaneously and frequency resources as current systems. Furthermore, this technology can manage challenges related to bandwidth limitations by exploiting the frequency resource concurrently while expecting no additional physical facilities. This paper studies the concern of boosting the coexistence of the three heterogeneous services across similar Radio Access Network (RAN) architecture whose specific characteristics and requirements will be elaborated later. The eMBB traffic can be analyzed to be a direct extension of the 4G broadband service. Massive payloads illustrate it, and the device operations stay stable over a prolonged period. That permits

the network to manage wireless resources to the eMBB devices ensure that no more than two devices using the same resource concurrently.

In this paper, we focus on URLLC in 5G wireless networks. Taking the above challenges into account, the key contributions of this work are summarized as follows:

1. The 5G main techniques (i.e., eMBB, mMTC, URLLC) are introduced, which act as different services in various 5G applications.
2. The requirements of the future application and the fundamental open problems toward the design of NOMA-based RAN are briefly explained.
3. An overview of the new generation of radio access technology is conducted, i.e., NOMA. Further, the basic principles of NOMA, including superposition coding (SC) and successive interference cancellation (SIC) for downlink channels, are presented.
4. Considering NOMA, we propose a novel algorithm where the downlink results are combined with D2D mode and utilize Greedy Asynchronous Distributed Interference Avoidance Algorithm (GADIA) to establish two different application scenarios. Mainly, we analyze the max–min fairness optimization problem for a given fairness index in the network. The proposed optimization framework has been proven theoretically and confirmed numerically by comparing with Monte Carlo simulations for the same system model as the one introduced in [15].
    To the best of our knowledge, NOMA-based D2D is a device group on the max–min fairness optimization problem that was not yet studied in the context of the heterogeneous network where both the cellular and D2D devices are active.

The remainder of this paper is organized as follows. In Section 2, we overview the related work. In Section 3, we consider the 5G network and explain the challenges and research directions. The main principles of NOMA and performance analysis are provided in Section 4. Sections 5 and 6 formulates and solves the max–min optimization problem subject to the power constraints. Finally, the main conclusions and future work are given in Section 7.
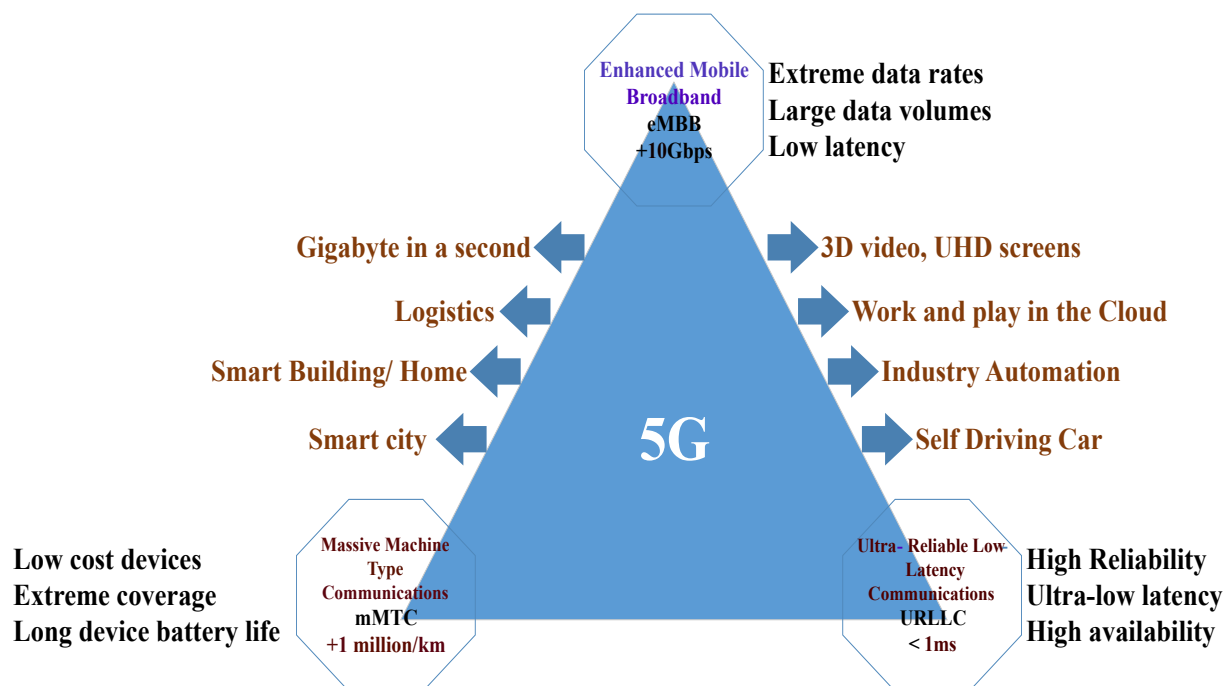
Figure 1. Features and improvement plan for future 5G.

2. Related Work

Most general 5G deployments heterogeneous services are permitted to coexist within the same network; moreover, the authors in [16] investigated the network structure slicing based on 5G. Accordingly, recent research works of 5G architectures concentrate on designing a network with low-cost base stations, low latency and high bandwidth abilities to sustain the diverse 5G applications. That introduces the front link bandwidth requirement exceeding what the current technologies, for instance, Common Public Radio Interface (CPRI), can provide the necessity for new transport technologies [17,18]. A promising method to perform these design purposes is by employing the concept of function split that is specified by 3GPP standard [19], and it has been comprehensively studied in [20]. The 5G slicing network architecture in the existence of orthogonal slicing has been introduced in [16,21]. The dynamic scheduling of the downlink multiplexing for URLLC traffic over continuing eMBB transmissions is investigated in [5,22] by abstracting the procedure at the physical layer. The authors in [23] propose a cost analysis of some of the PHY layer function split prospects. Lien et al. in [24] consider the role of feedback in the regard of resource allocation for mMTC and URLLC in a new radio (NR) setting. Chitimalla et al. in [25] proposed a scheduling policy to decrease jitter. In [26], Mahmood et al. investigated the orthogonal resource allocation for mMTC and eMBB users through imputation the intercell interference. Grant-free uplink transmissions are deemed in [27] for the three services by evaluating a particular transmission/modulation/spreading methods for sustaining the three services. In contrast, numerous works present an extensive analysis of the impact that virtualization has on latency, furthermore testing the performance through various virtualization circumstances, e.g., kernel-based virtual machine (KVM) and VirtualBox investigated in [28–31]. Additionally, authors in [32–36] have already briefly described Random-access (RA) based schemes for the URLLC and mMTC [35,36] individually. A comprehensive review of the scheduler prospects in 5G NR is provided in [3], particularly classifications of new scheduling formats and levels of freedom added to promote URLLC and additional services. Moreover, several studies on the dynamic link adaptation and MAC layer for URLLC have been conducted in the literature (see, e.g., [37,38]). The fundamental concept of diversity and coding techniques were investigated in [39]. The hybrid automatic repeat request (HARQ) enhancements discussed in [40,41], and the authors in [42,43] addressed the variable transmission time intervals (TTIs). Slalmi et al. in [44] studied the optimization of the efficient bandwidth allocation for each device using 5G technology in orthogonal frequency-division multiplexing (OFDM) systems and reported an improvement for reliable wireless communications. NOMA concept is also introduced in conjunction with other techniques, such as MIMO or mmWave [45,46]. Many previous works have investigated NOMA technology's potential contributions to 5G-and-beyond wireless networks [47–49]. However, these works are built on traditional wireless techniques driven by broadband communications and are often inefficient for URLLC.

3. 5G Overview

There is a massive increase in the market for wireless services demanding faster, higher-capacity networks owing to numerous devices that require to be connected to the Internet simultaneously, for instance: mobile phones, computers, tablets, smartwatches, etc. The IoT applications earn significant attention that is employed with restriction to the high-speed broadband connection. For instance, E-Health and E-Transportation fuel a need for massive connectivity and URLLC. These devices constantly require higher reliability communications within low latency constraints. The 5G NR consider bringing provision for diverse applications with an extensive range of communication requirements. These requirements are linked with three generic service classes, known as eMBB, mMTC, and URLLC. The 3GPP has considered these service classes as the primary leading drivers for designing 5G. The major steps in the evolution of telecommunication technologies are summarized in Table 1.

Table 1. The evolution of telecommunication technology.

Generation	The Key Technologies	Services
1G	AMPS	Analog voice service
2G	GSM, TDMA, CDMA	Digital voice service
3G	WCDMA, CDMA2000, TD-SCDMA, WiMAX	Simultaneous transmission of voice and data
4G	TD-LTE, FDD-LTE	Fast transmission of voice, data, video, image.
5G	Millimeter Wave, Massive MIMO, Micro Base Station, D2D, Beamforming, NOMA	HD video, smart home, etc

3.1. Enhanced Mobile Broadband (eMBB)

eMBB is a communication service that enables high data rate communications, which is crucial for numerous human-oriented applications, such as video streaming, high-traffic mobile broadband service, virtual reliability. Moreover, online gaming and the scenarios, including 3D/Ultra HD video live broadcasts that can be shared and accessed anytime and anywhere, dramatically increase wireless traffic. Indeed, the massive demand for an extended data rate was the primary motivation for the earlier evolution of the cellular systems' generations, including 3G and 4G. 5G has certain advantages in large bandwidth and low latency requirements. It intends to offer a peak data rate of 20 Gbps and assure a decent data rate (50–100 Mbps) of ubiquity adverse provisions, e.g., in a dense area under high mobility. Additionally, the target for user data traffic latency is 4 ms.

3.2. Massive Machine Type Communication (mMTC)

mMTC indicates the communication service for connecting a massive number of devices. This service can foster the fulfilment of the IoT by implementing effective Internet connectivity. Extensive connectivity is expected for numerous applications associated with environmental measurements, traffic monitoring, asset tracking, resource utility management, and the smart grid. These applications may require further communication requirements besides extensive connectivity. For example, many devices are battery-powered and should stand for a long time, e.g., a couple of years, without replacing them. For instance, smart meters might be used in locations with high penetration loss. Therefore, they need prolonged network coverage for their services. Furthermore, a low-cost transceiver is fundamental for recognizing various IoT applications. Some of these characteristics was partially discussed in Long Term Evolution (LTE) releases 13 and 14 by proposing new UE categories, such as UE CAT-0, Cat-M1, and Narrowband Internet of Things (NB-IoT) [PI] [44].

3.3. URLLC

URLLC is a new communication service that is recognized for cellular systems. This service is crucial for sustaining numerous emerging mission-critical applications, mainly industrial automation, vehicular safety communications, and remote surgery. URLLC has rigorous requirements on availability, latency, and reliability [45]. This implies that the communication link should be available most of the time to transfer data in a short period and with high reliability. The 3GPP intended to provide URLLC for short data payloads (up to 32 Bytes) with the user plane latency of 1 ms (measured at layer 2 and layer 3) and the reliability of $1-10^{-5}$ (99.999%). An asymmetric latency budget of 0.5 ms is defined for both up-link and down-link to assure the latency target. The information must be carried at the higher layers quickly and error-free with a significant level of integrity, reliability, and confidentiality [46]. As long as these specifications are adequate for many mission-critical applications, some other applications have more precise requirements and further demands. Therefore, the 3GPP may define more rigorous requirements for URLLC service in later 5G releases.

3.4. Challenges and Research Directions

NOMA-based RAN affords a powerful and promising platform for achieving future application requirements. Various essential issues persist in thoroughly assessing the gains that can potentially endure from such an architectural integration. The following Table 2 addresses the fundamental open problems toward the design of NOMA-based RAN.

Table 2. Integration of NOMA in different radio access network technologies.

NOMA Use Case	Opportunities	Challenges and Research Directions
eMBB	Downlink NOMA, D2D NOMA, edge caching, Better spectral efficiency and user's fairness, Increased number of several users satisfying their QoS	Design a comprehensive RRA, IM, caching placement and strategy that includes down-link NOMA and D2D NOMA
mMTC	Grant-free NOMA in the uplink transmissions, Reduced energy consumed in signaling, Higher transmission success probabilities compared to OMA	Optimize the number of multiplexed users through NOMA
URLLC	Dowanlink NOMA by increasing the networking availability, Up link NOMA by reduce the collision, Data storage and computation at the network edge	Analyze the joint provision of high reliability and low latency

4. Principle of NOMA System

In this section, we briefly explain the fundamental principles of orthogonal multiple access (OMA) and NOMA. Traditional OMA schemes are: frequency-division multiple access (FDMA), time-division multiple access (TDMA), code-division multiple access (CDMA) and orthogonal frequency-division multiple access (OFDMA) used for 1G, 2G, 3G, and 4G, respectively. Multiple users are allocated to orthogonal radio resources in frequency, time, code-domain, or combinations. More particularly, every user sends a unique user-specific signal through its unique frequency resource in FDMA. Therefore the receiver can quickly identify all users' data in their identical frequency bands, respectively. Furthermore, in TDMA, a particular time slot is allocated to each user; accordingly, it is straightforward to notice the distinct users' signals at the receivers in the time domain. In CDMA, various users can share the same time-frequency resources, while the transferred symbols of distinctive users may be correlated to orthogonal spreading sequences, e.g., Walsh-Hadamard codes. Hence, a low-complexity decorrelation receiver enables the use of multi-user detection (MUD). OFDMA can be considered an intelligent combination of FDMA and TDMA where the radio resources are orthogonality divided in the time-frequency grid [47]. Theoretically, as an advantage of orthogonal resource allocation, there is no interference between OMA systems users. Consequently, low-complexity indicators with linear complexity can assign different users' signals. The maximum number of sustainable users is tightly limited by the available number of orthogonal resources in conventional OMA schemes which results in reducing the extensive connectivity in 5G. Moreover, it has been proven that OMA cannot regularly obtain the maximum achievable sum-rate of multi-user wireless systems. Unlike OMA techniques, NOMA can accomplish

the multi-user capacity with the aid of time-sharing or rate-splitting if required [48]. To address the above constraint of OMA, NOMA has been recently introduced as a potential solution. The key characteristic of NOMA scheme is to maintain reliable communication for a massive number of users higher than the number of orthogonal resource slots supported by non-orthogonal resource allocation [49]. The complex inter-user interference cancellation can perform at the cost of an enhanced receiver complexity, e.g., computational complexity with polynomial or exponential order. An effective way to eliminate the interference is to exploit powerful multiuser detectors. NOMA schemes can be introduced in two domains: power-domain NOMA and code-domain NOMA. NOMA in the power-domain, for distinct users, indicates different power levels based on their channel quality, as long as the same time-frequency-code resources are shared between multiple users. NOMA power-domain utilises the users' power-difference to identify various users concerning successive interference cancellation (SIC) at the receiver side. The NOMA Code-domain is comparable to CDMA or multi-carrier CDMA (MC-CDMA) unless its decision to employ the low-density sequences or non-orthogonal sequences has a low cross-correlation [50].

4.1. SIC Technology

The basic principle of SIC technology is to adopt a step-by-step interference elimination strategy. In the received signal, users are judged one by one according to the order of power. After the amplitude is restored, the signal's multiple access interference are subtracted from the received signal. Then, judge the remaining users again, and repeat the operation until all the multiple access interference is eliminated. Therefore, SIC technology has more significant advantages for users with weak signals.

Figure 2 illustrates the typical diagram of SIC strategy for three users. Assume the three users superimposed on the same sub-channel are x_1, x_2, x_3 where the corresponding channel gains are, respectively, h_1, h_2, and h_3 under $h_1 > h_2 > h_3$ condition; i.e., the users are ranked based on SNR. When the SIC receives the signal from the wireless channel, it first sorts all user signals according to their power from strong to weak and then performs detection according to the power. Since x_3 has the strongest signal power and the corresponding SINR, SIC detects x_3 first, and then detects x_2 after obtaining the signal of x_3, and finally x_1.
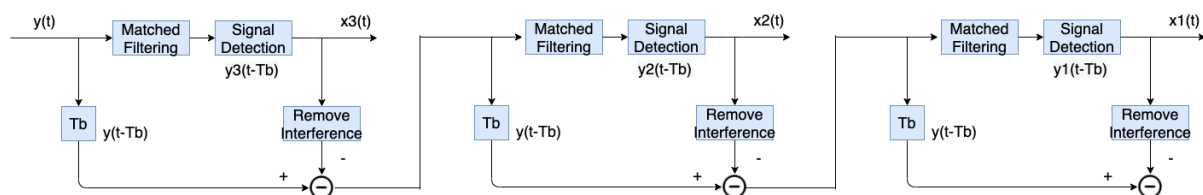

Figure 2. SIC receiver principle.

There are several potential problems with SIC receivers that may adversely affect the performance of NOMA. The first disadvantage is that it is more sensitive to the power level, i.e., if two or more UEs have similar power, the performance will be significantly reduced. If two or more UEs have similar channel gains, the NOMA scheme will assign them the same power allocation coefficient, lowering the SIC's performance. Another drawback of SIC is its high dependency on the correct decoding of the first UE. If the first UE fails to decode successfully, this decoding error propagates through the lower-order of UEs and results in system performance degradation. Another limitation of the SIC receiver comes from the decoding delay for low-order UEs due to the nature of the SIC serial decoding process. To reduce the decoding delay, NOMA must group a small number of UEs with significantly different channel gains, limiting the system capacity. Finally, because the NOMA protocol serves users based on the SIC sequence, the rate achieved between users is different.
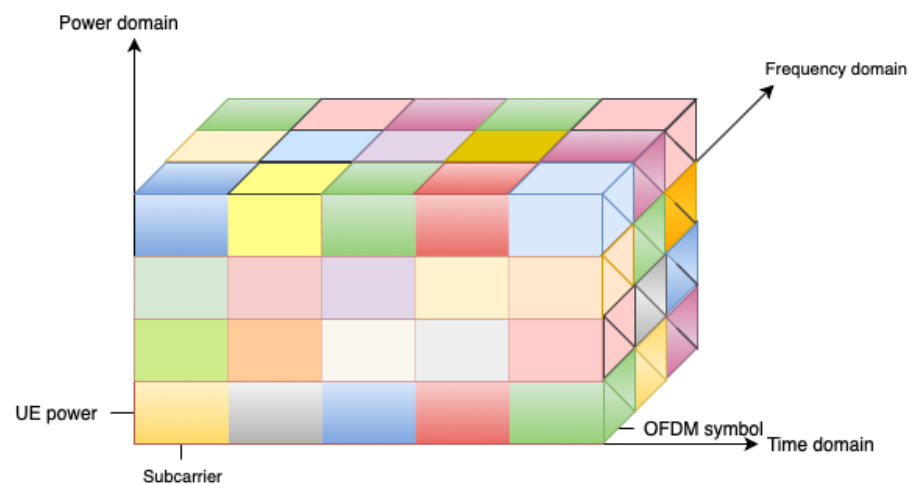
4.2. Resource Allocation

In the conventional OFDMA system, just a single symbol is mapped to a sub-carrier. Distinct characters are sent on different sub-carriers that are orthogonal and do not interfere with each other. Accordingly, the total number of transferred symbols is limited by the number of orthogonal sub-carriers. Figure 3 illustrates a typical diagram of resource allocation carried by user information on the frequency domain and time domain planes of OFDM technology. The smallest unit in the frequency domain represents a sub-carrier with a specific bandwidth. The smallest unit in the time domain can be considered a particular length of the OFDM symbol. Each resource unit is a resource plane composed of one subcarrier and one OFDM symbol. Since each resource unit's user power is the same, the OFDM technology did not develop a power domain in terms of the power domain. It is a two-dimensional planar resource carrying system, and its wireless channel transmission capacity is relatively weak.
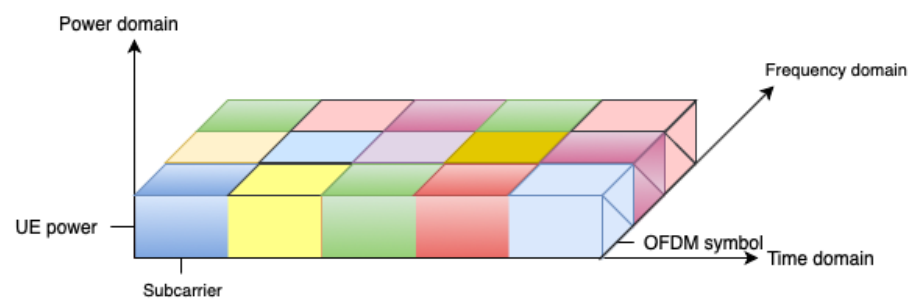
Figure 3. Resource allocation of OFDM technology.

NOMA can be obtained in a recently emerged new domain, particularly in the power domain. In contrast to the multiple access schemes based on time, frequency, code-domain, or combinations as illustrates in Figure 4. Figure 4 illustrates a diagram of resource allocation in NOMA scheme. The corresponding frequency domain, time domain plane, and OFDM resource plane are the same. The difference is that it has developed the power domain application. Unlike OMA technology that can only allocate a unit resource (i.e., time/frequency) to one user, NOMA can take advantage of allowing different power and assign a unit resource to numerous users.

Figure 4. Resource allocation of NOMA technology.

NOMA technology is a three-dimensional resource bearing system, so wireless channels' transmission capacity is much stronger than OFDM. Distinctive signals formed by various users are instantly superimposed on each other after classic channel coding and modulation at the transmitter. Multiple users share the same time-frequency resources, fol-

lowed by exposure to the receivers by MUD algorithms such as SIC. Accordingly, the SE is magnified at the cost of an enhanced receiver complexity compared to conventional OMA. Furthermore, in the information-theoretic literature, it has been broadly acknowledged that non-orthogonal multiplexing utilizing superposition coding at the transmitter and SIC at the receiver outperforms classic orthogonal multiplexing.

Figure 5 illustrates the downlink transmitter and receiver of the NOMA system management process per sub-channel. The system consists of a single base station (BS) and multiple users. The BS transmits the signal to the user set $u \in 1, 2, \ldots, U$ through the sub-channel set $v \in 1, 2, \ldots, V$, and all receivers are equipped with SIC per each sub-channel and single antenna. Assuming that the system bandwidth is BW, each sub-channel bandwidth is $B_v = BW/V$. Assume P_{BS} denote the transmit power of the BS. Furthermore, let P_v and $P_{u,v}$, respectively, represent the power of the sub-channel and the power of the uth user on the sub-channel v.

As an example, assume that the base station serves three users denoted as $UE1$, $UE2$, and $UE3$ where their channel responses are, respectively, h_1, h_2, and h_3, and their corresponding signal-to-noise ratios are 20 db, 10 db, and 0 db, respectively. As it can be readily checked, the h_1 channel condition is the best one with the highest signal-to-noise ratio. The h_2 channel condition is worse than h_1, and the h_3 channel condition is the worst. Next, we will analyze the signal processing situation at the transceiver end the principle of NOMA. A user over strongest channel should not remove interference from others. The best user decodes the signal containing possible interference and shares the decoded signal with the other users. The user over the second strongest channel decodes its signal by removing the best user's interference and shares it with the others. The last user over the worst channel will have zero interference from other users after applying (n−1) SIC and decodes its signal.
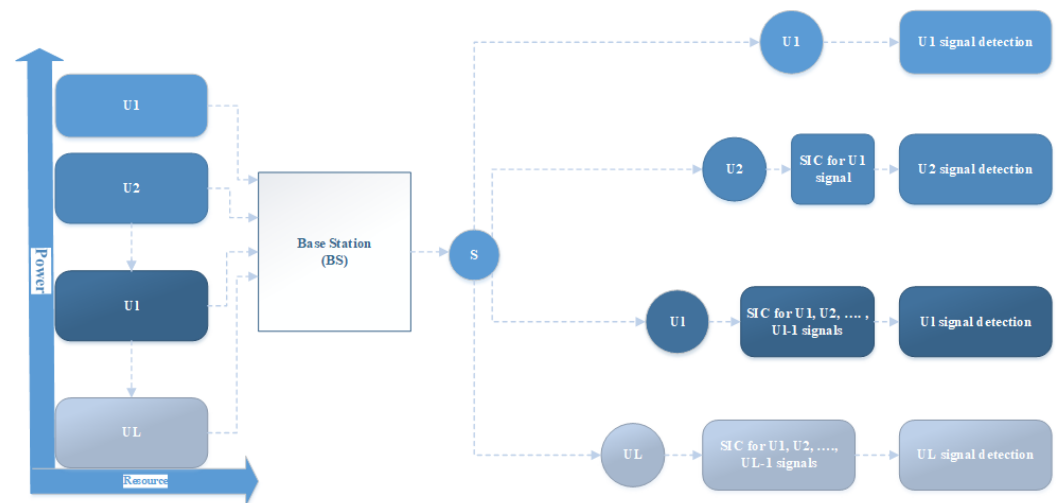

Figure 5. Downlink sub-channel system model of NOMA.

Assuming that n users are assigned for each sub-channel, the signal sent by the BS to each sub-channel can be expressed as:

$$x_v = \sum_{i=1}^{n} \sqrt{P_{i,v}} s_i \qquad (1)$$

where s_i is the modulated symbol of the ith user. According to the principle of NOMA, multiple users can be superposed on the same sub-channel where each user not only receives its own signal but also receives possible interference signals from other users on the same channel. The received signal at the jth user over vth sub-channel can be expressed as:

$$y_{j,v} = x_v h_{j,v} + w_{j,v}$$
$$= (\sum_{i=1}^{n} \sqrt{P_{i,v}} s_i) h_{j,v} + w_{j,v}$$
$$= \sqrt{P_{i,v}} s_i h_{j,v} + \sum_{i=1, i \neq j}^{n} \sqrt{P_{i,v}} s_i h_{j,v} + w_{j,v} \tag{2}$$

where $h_{j,v} = g_{j,v} PL^{-1}(d)$ is the coefficient of sub-channel v from BS to the jth user, $g_{j,v}$ is the Rayleigh fading channel gain and $PL^{-1}(d)$ is the path loss function between the *BS* and the jth user at distance d. Further, $w_{j,v}$ is the additive white Gaussian noise (AWGN) with zero mean and variance σ_v^2 where $\sigma_v^2 = E[|w_{j,v}|^2]$.

Then the signal to interference plus noise ratio (SINR) (without SIC) received by the jth user can be expressed as [51]:

$$SINR_{j,v} = \frac{P_{j,v}|h_{j,v}|^2}{\sigma_v^2 + \sum_{i=1, i \neq j}^{n} P_{i,v}|h_{j,v}|^2} \tag{3}$$

The normalized signal to interference plus noise ratio can be expressed as [51]:

$$SINR_{j,v} = \frac{P_{j,v} H_{j,v}}{1 + \sum_{i=1, i \neq j}^{n} P_{i,v} H_{j,v}} \tag{4}$$

where $H_{j,v} = \frac{|h_{j,v}|^2}{\sigma^2}$ represents the channel response normalized by noise of the jth user on sub-channel v. Recalling the principle of SIC approach and the optimal decoding order, we assume that the channel conditions of n users superposition on sub-channel v are sorted by the channel response normalized by noise (*CRNN*) as:

$$|H_{1,v}| \geq |H_{2,v}| \geq, \ldots, \geq |H_{j,v}| \geq |H_{j+1,v}| \geq, \ldots, \geq |H_{n,v}| \tag{5}$$

As previously mention, in NOMA protocol, more power is allocated to user with poor channel conditions. As a result, the power allocated to each user on sub-channel v can be expressed as:

$$|P_{1,v}| \leq |P_{2,v}| \leq, \ldots, \leq |P_{j,v}|, \leq |P_{j+1,v}| \leq \ldots, \leq |P_{n,v}| \tag{6}$$

After removing the interference from the users having poorer channel condition, the received signal to interference plus noise ratio (with SIC) of user jth can be expressed as:

$$\widetilde{SINR}_{j,v} = \frac{P_{j,v} H_{j,v}}{1 + \sum_{i=1}^{j-1} P_{i,v} H_{j,v}} \tag{7}$$

Recalling the Shannon's capacity expression, the data rate of the jth user on sub-channel v is given by:

$$R_{j,v} = B_v log(1 + \frac{P_{j,v} H_{j,v}}{1 + \sum_{i=1}^{j-1} P_{i,v} H_{j,v}}) \tag{8}$$

The sum rate of the sub-channel v is given by:

$$R_v = \sum_{j=1}^{n} B_v log(1 + \frac{P_{j,v} H_{j,v}}{1 + \sum_{i=1}^{j-1} P_{i,v} H_{j,v}}) \tag{9}$$

Therefore, the overall sum rate of system can be given by:

$$R = \sum_{v=1}^{V} \sum_{j=1}^{n} R_{j,v} \tag{10}$$

4.3. Principle of Transceiver

Figure 6 illustrates a typical scheme for the power allocation, the uplink and downlink of the NOMA system management process per sub-channel. Multiple users send their signals to the BS in the same Resource Block (RB) in uplink transmission, as illustrated in Figure 6b. The BS identifies all the users' messages with the assistance of SIC. The main differences between uplink NOMA and downlink NOMA can be summarized as follows:

- **Transmit Power:** in regards to downlink NOMA, the users' transmit power in uplink NOMA does not have to be different. It is based on the channel conditions of each user. If the users' channel conditions are considerably different, their received SINR can be somewhat different at the BS, irrespective of their transmit Power.
- **SIC Operations:** the uplink NOMA and downlink NOMA users' SIC operations and interference are also somewhat different. In particular, as displayed in Figure 6c for downlink NOMA, the signal of user n is disinfected from the interference caused by user m. In precise, that is performed by first detecting the more powerful signal of user m, modulating it and then deducting it from the composite signal. That implies SIC operation is carried out on a strong user in the downlink for cancelling the weak user's interference. By contrast, SIC is performed at the BS in uplink NOMA to distinguish the strong user n first by treating User m as interference, as presented in Figure 6b. Then it demodulates the recovered signal and subtracts the interference imposed by user n to identify user m.
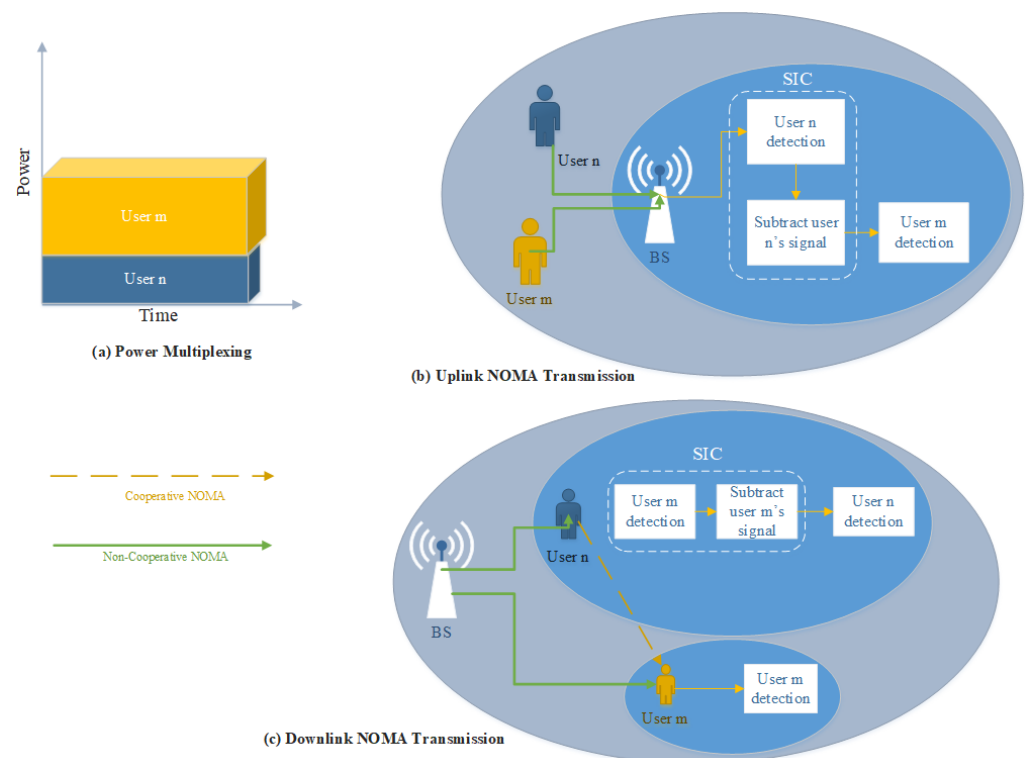

Figure 6. Principle of transceiver of NOMA system.

5. Proposed Algorithm

NOMA technology is considered as one of the key technologies of the 5G mobile communications [52]. It is able to distribute the power and SIC technology allows multiple users to share the same resource, such as time, frequency rate, etc. As a result, it is expected to improve the system performance in terms of throughput and energy efficiency. D2D and cooperative relaying technology are the key technologies in 5G systems. There is also a surge of interest in utilizing NOMA technology combined with D2D and relay technology. The combination of NOMA and D2D gives full play to the advantages of the two hotspot technologies and is very suitable for 5G scenarios. An interesting algorithm

called GADIA is proposed in the literature [53]. The algorithm itself provides a simple fully distributed dynamic frequency allocation strategy, without any information exchange between autonomous systems equipment and no knowledge about the existence of other equipment autonomous entity. Moreover, it can be used as a part in conjunction with any reasonable radio channel model. In the GADIA calculation, each node has the knowledge of possible interference that it may encounter, and then selects the spectrum band with the least interference. In this work, we use the combination of NOMA and D2D to build a system model, and employ the GADIA algorithm to allocate D2D strategies.

5.1. System Model

We consider one BS and K cellular users as $UE_1, UE_2, \ldots, UE_K$. Among them, UE_1 is the closest to the base station with the best channel conditions, and UE_K is at the edge of the cell with the worst channel conditions. In the two-stage NOMA-D2D model, UE_2 can act as a relay for UE_1, while UE_2 can directly communicate with UE_3 as D2D transmitter. In the first stage, the BS sends signals to UE_1 and UE_2 in NOMA mode; in the second stage, UE_2 acts as a relay and D2D transmitter and also sends signals to other users in GADIA mode.

5.2. Algorithm Design

The transmitted signal by the BS can be written as:

$$x(t) = \sum_{k=1}^{K} x_k(t) \sqrt{\alpha_k P_T} \tag{11}$$

where P_T denotes the base station transmit power and α_k is the power allocation factor for the UE_K power allocation factor. The power allocated to each UE_k can be written as:

$$P_k = \alpha_k P_T \tag{12}$$

and the received signal at the UE_k can be expressed as:

$$y_k(t) = h_{bu} \sum_{k=1}^{K} x_k(t) \sqrt{\alpha_k P_T} + n_{uk} \tag{13}$$

where n_{uk} represents the received Gaussian white noise, and h_{bu} indicates the channel from the base station to UE_K. UE_1 first detects $x_k(t)$ and eliminates the interference of the signal to $y_k(t)$, and then detects $x_{k-1}(t)$ and eliminates the interference of the signal to $y_k(t)$, until x_1 is checked. Therefore, the SINR for UE_K can be written as:

$$SINR_k = \frac{P_k |H_{bu}|^2}{\sum_{i=1}^{K-1} P_i |H_{bu}|^2 + 1} \tag{14}$$

Using NOMA, the throughput of each UE in bps can be expressed as:

$$R_k = W log_2 \left[1 + \frac{P_k |H_{bu}|^2}{\sum_{i=1}^{k-1} P_i |H_{bu}|^2 + n} \right] \tag{15}$$

where W denotes the transmission bandwidth. Furthermore, let J denote the fairness index which can be written as:

$$J = \frac{(\sum R_k)^2}{K \sum R_k^2}. \tag{16}$$

In Equation (16) represents the fairness of sharing system capacity among UEs. When J is close to 1, the capacity of each UE is close to each other. We can set the goal of the

power allocation mechanism to maximize the total capacity of the NOMA system under fairness constraints. The nonlinear optimization problem can be expressed as:

$$
\begin{aligned}
\max_{\alpha_k} \quad & W \log_2 \left[1 + \frac{P_k \, |h_{bu}|^2}{\sum_{i=1}^{k-1} P_i \, |h_{bu}|^2 + n} \right] \\
\text{s.t.} \quad & J \geq J' \\
& P_k \geq 0, \forall k \\
& \sum_{k=1}^{k} \alpha_k \leq 1 \\
& \sum_{k=1}^{k} P_k \leq P_T
\end{aligned}
\tag{17}
$$

where J' is the target fairness index in the network.

5.3. Energy Model

In this section, we analyze the energy efficiency (EE) of NOMA systems. Considering the downlink, the total power consumed P_{tot} by the BS can be written as:

$$
P_{tot} = P_s + P_T \tag{18}
$$

In (18), P_T is the power consumed by the signal transmission, and P_s is the power consumed due to other hardware and signal processing operations. Therefore, the total transmission power can be expressed as the sum of the energy consumed by the signal transmission power and the energy consumed. EE can be defined as the ratio of the throughput over the total power. Mathematically speaking, it is given as [42].

$$
EE = \frac{R_t}{P_{tot}} \tag{19}
$$

6. Results and Discussion

In this section, we demonstrate the performance of the NOMA scheme under consideration. First, we study the boundary rate comparison of OFMA and NOMA. Next, the performance of EE and SE is investigated. Finally, we consider a D2D system model for two distinct condition (i.e., rural and urban) and investigate the behaviour of GADIA algorithm. Figures 7–9 illustrate the main steps of our simulation process; i.e., random pairing algorithm, channel state sorting-pairing algorithm and orthogonal pairing algorithm. The simplest grouping method is the Random Pairing Algorithm (RPA) [54]. For grouping, the BS chooses users at random from the user pool. The algorithm has a Committed Information Rate when CIR0 is the rate threshold users establish based on their business needs. Users that meet the rate can be chosen at random by the BS. However, the system performance is poor because this method does not consider the user's channel state. The Channel State Sorting–Paring Algorithm (CSS-PA), on the other hand, considers that there are N users to be assigned. First, organize the candidate set's channel conditions in ascending or descending order. If the number of users is even, divide them in half and divide them into groups A and B. When the number of users is odd, the first $\frac{N}{2}$ users are in group A, the $(\frac{N}{2} + 1)$th user is in group B, and the rest of the users are in group C. After the grouping, match pairwise in sequence A users and B users [55]. Moreover, the OPA algorithm allows users with the best channel orthogonality to pair [56].In turn, match each user in the candidate user set. Then, calculate the orthogonality factor for each match, and choose the user with the highest orthogonality factor. Therefore, delete the matched users from the group of candidate users.
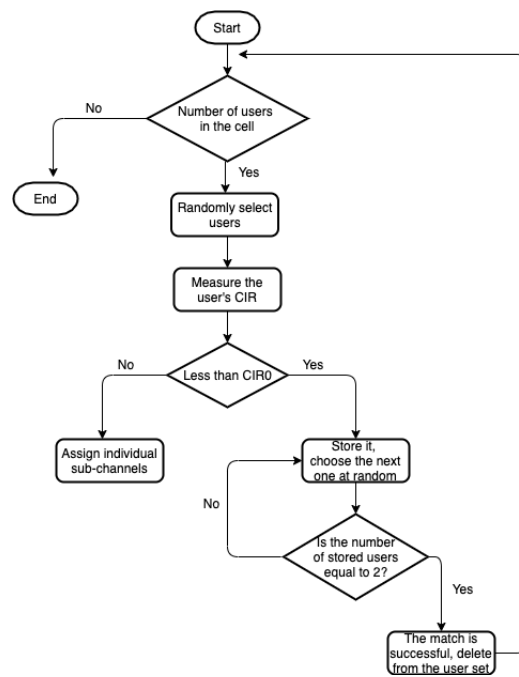
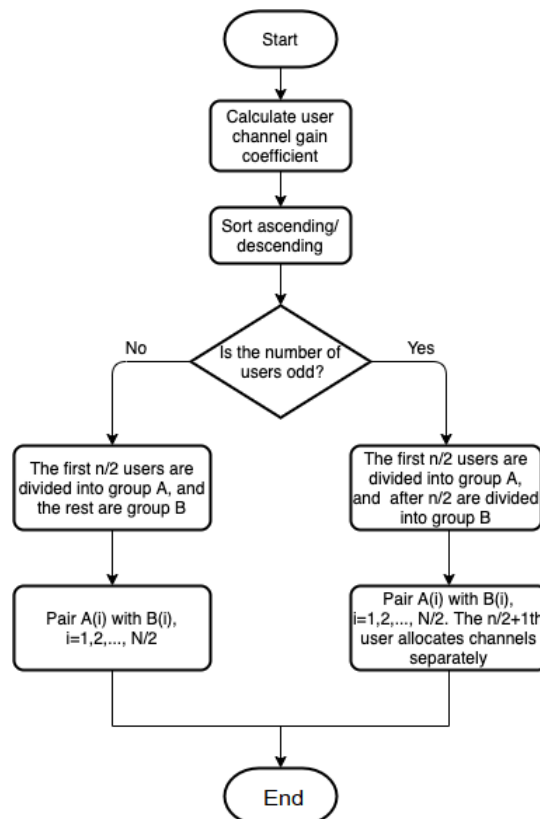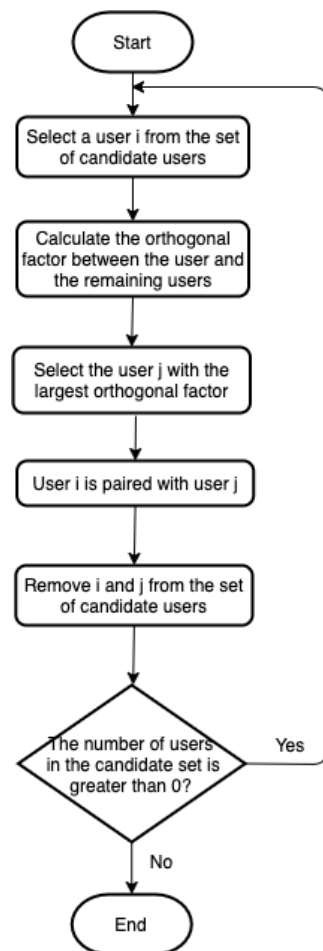**Figure 7.** The flowchart of random pairing algorithm.

Figure 8. The flowchart of channel state sorting-pairing algorithm.

Figure 9. The flowchart of orthogonal pairing algorithm.
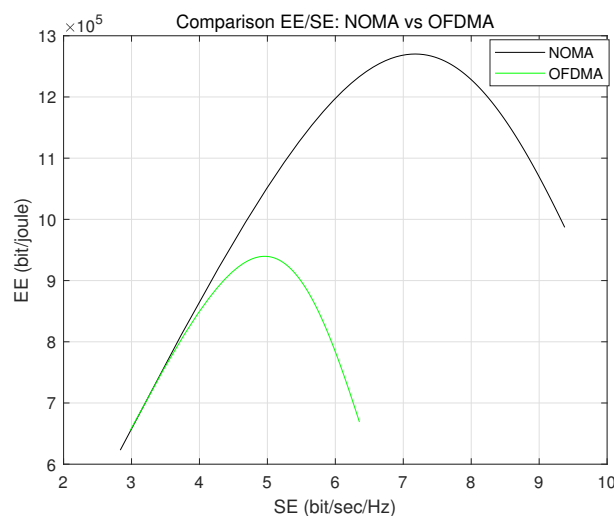
6.1. Comparison NOMA vs. OFDMA

Figure 10 illustrates the boundary rate comparison of OFMA and NOMA. To better illustrate this, a pair of users communicating with a BS over an AWGN channel is considered as an example. As it can be observed from Figure 10, utilizing NOMA leads to a larger capacity region compared to OFDMA. The rate of user 2 with NOMA is about 3 times higher than the one with OFDMA. That is regarded as a significant superiority of NOMA over OFDMA. Intuitively, the users associated with small normalized channel gains should be allocated higher power levels to improve their received SINR.More particularly, the users with higher channel gain require less power and they can successfully detect the data with a high probability as a benefit of SIC. In the absence of preventative measures, the SIC induced error propagation may have a severe effect on the error probability as the number of users increase. Based on this optimal SIC-detection order, any user can detect its information without substantial interference-contamination imposed on the other users whose normalized channel gain is smaller than that of this user.
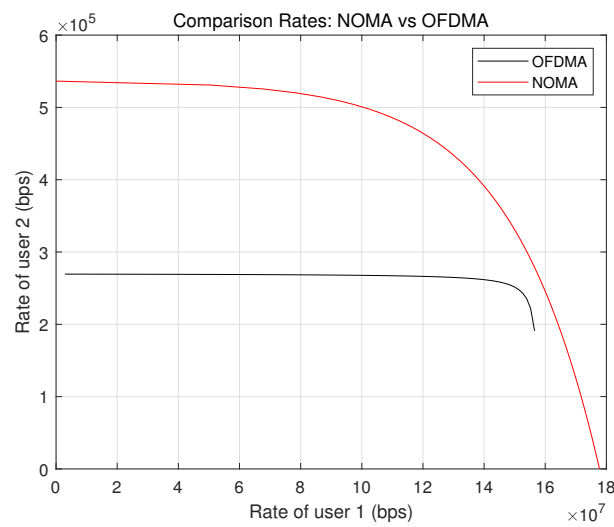
Figure 10. The boundary rate between NOMA and OFDMA for downlink.

In Figure 11, we investigate the performance of EE and SE. We assume the power consumed due to hardware and signal processing operations is $Ps = 100$ W. For Equation (19) the EE peak can be assumed in NOMA is 3696.4 bit/joule, and OFDMA's peak is 4.5 bit/joule. The result of NOMA is much better 370 times than the result of OFDMA. Furthermore, the NOMA scheme has the potential of striking a more attractive trade-off between SE and user-fairness. However, some advanced user pairing and power allocation methods and powerful channel coding schemes reduce the error probability. Furthermore, NOMA's down-link could achieve the capacity region, while OFDMA is suboptimal in general. However, rate-fairness is not maintained at this optimal point since the low-power user rate is much lower than that of the higher-power user when the distinction between the two users' received powers is high. Note that the simple two-user case results can be extended to the general case of an arbitrary number of users. In this way, the user with the most substantial normalized channel gain can cancel interference from all the other users and thus is least contaminated. Then, simulations in Figure 12 compare the sum data rate in a rural area at a given EE and SE when using NOMA versus OFDMA. The simulation were performed in MATLAB, and the results showed that NOMA achieves a better performance than OFDMA in the categories compared.

Figure 11. The energy efficiency and spectral efficiency comparison for NOMA and OFDMA.
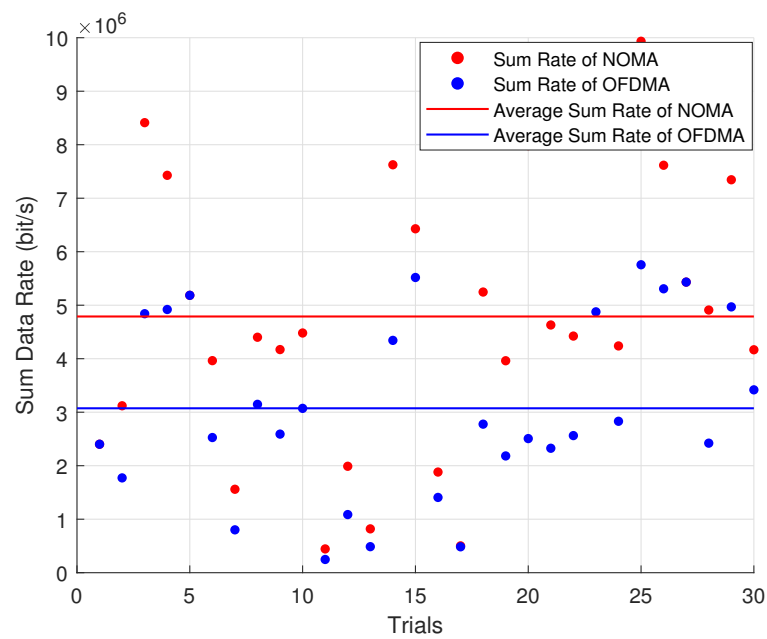
Figure 12. Data rate comparison of NOMA vs. OFDMA in rural areas.

6.2. D2D System Model

In this scenario, we consider a D2D system model containing a cellular system where the number of assigned users follow a Poison Point Process (PPP) and the user's spatial position is randomly distributed following a uniform distribution. We assume the D2D users are also distributed following a Poisson point process. In this part, the downlink data of the base station received by the UEs receiving end is optimized in combination with D2D. For our study, we consider a square cell, where the BS is omnidirectional and placed at the cell centre. The research is focused on the downlink scenario. A crucial part of the downlink is power control. Furthermore, we reflect shadow fading, represented as a Gaussian random variable X with zero mean and standard deviation in dB. Recalling the GADIA algorithm, we create two different scenarios where GADIA allocates the D2D users in rural, and urban environment. For the convenience of the reader, the system parameters are summarized in Table 3 unless otherwise stated.

Table 3. System parameters.

Type	Urban	Rural
BS Sensitivity (dBm)	-100	-100
BS Coverage Area	1 km^2	1 km^2
UEs Max Tx Power	30 dBm	30 dBm
Path Loss Exponent	6	2

The GADIA algorithm is applied to a single random user deployed in the square unit. It then conducts a continuous iterative analysis of its performance, focusing on the interference of its behaviour. After several iterations, the GADIA algorithm is applied to different random users to obtain the average value of its performance. This process is iterated until the number of users is exhausted. The BS is illustrated as a red point in the square cell centre, BS users as blue squares, and D2D connections as green triangles. Figures 13 and 14 , respectively, illustrate distribution of UEs within a range of 1 km using GADIA for rural and urban scenario under consideration. After the frequency allocation using GADIA, D2D users will stay within their frequency or change into the BS users. That shows how users are stationed in this simulation model for the two studied environments. Moreover, D2D users (represented with triangles) are colored with green (if they remain to

utilise the non-shared frequency band) or cyan (if they turn to the shared frequency band). The figures represent how the D2D users are allocated with one or another frequency according to the number of interference signal received from these maps.
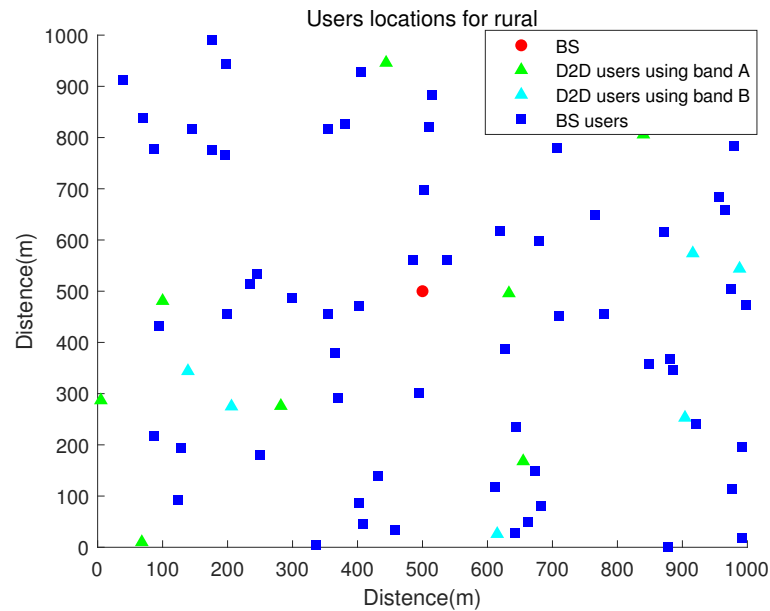

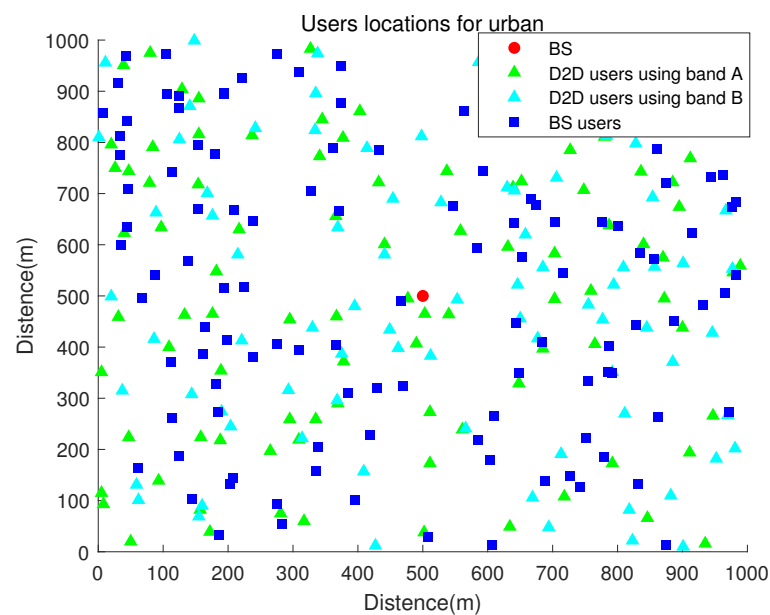**Figure 13.** Location map after GADIA for rural environment.

Figure 14. Location map after GADIA for urban environment.

After several iterations, the average interference of the two scenes stabilized, and the mean interference received by D2D users represent in Figures 15 and 16. The two other variables have been pursued along GADIA iterations to determine the whole network's performance ultimately: mean interference received by all network users (red line) and mean received interference of cellular network users (blue line). The simulation results for the two studied environments confirm how the D2D user's performance is enhanced at the expense of a deterioration of the cellular network users performance. The average interference of D2D users in cities is −60 dBm, and the moderate interference of D2D users in rural areas is −30 dBm. That effect can be described by increasing the interference that

cellular network users receive as the D2D users are allocated with the second frequency band (shared with the cellular network users).
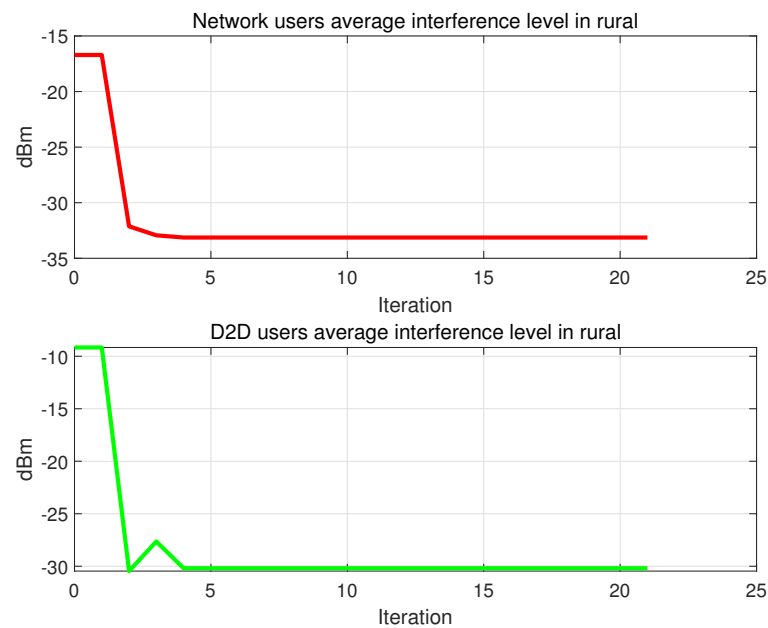

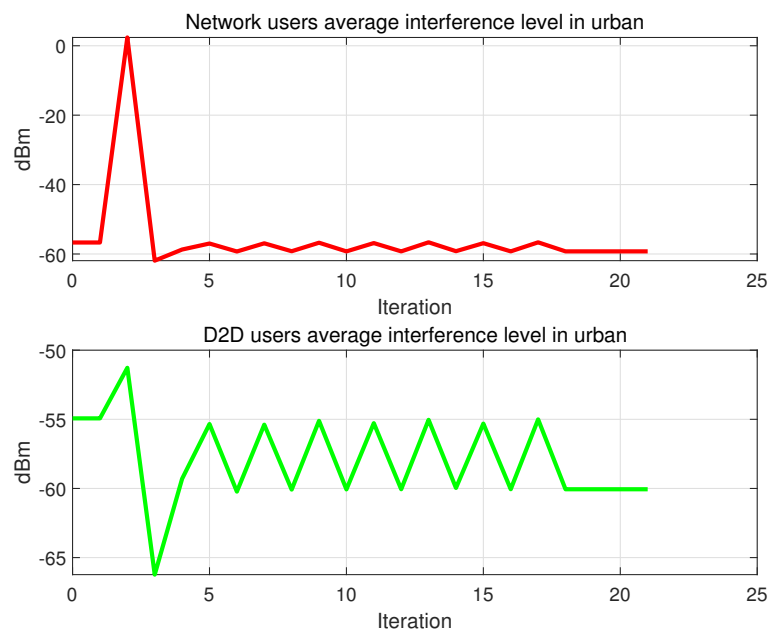**Figure 15.** Mean interference level in rural environment along GADIA iterations.

Figure 16. Mean interference level in urban environment along GADIA iterations.

As it can be seen in Figure 17, employing GADIA controlled frequency allocation brings improvements compared to random frequency allocation or not using any kind of frequency sharing approach. It should be noted that the final simulation result strongly relies on the initial deployment of users. For determining the actual behaviour of GADIA, it is required to run the simulation multiple times and analyze the mean performance afterwards. In this regard, we have repeated the experiments 200 times to obtain reliable results. To better emphasize this, we study the mean interference received by D2D users by applying three different approaches, i.e., GADIA controlled frequency allocation, random frequency allocation, and without using any frequency sharing. As it can be readily checked from Figure 17, the average interference for GADIA-based algorithm for urban scenarios is about −74 dBm which is much smaller than other resource allocation schemes.
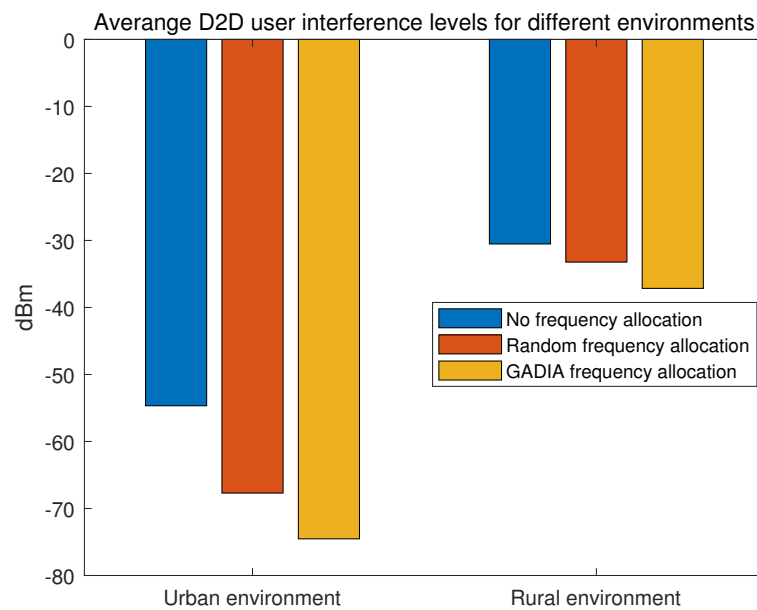
Figure 17. Mean D2D user interference level for different environments after 200 iterations.

7. Conclusions and Future work

Several innovative techniques have been introduced for the next mobile generation. The next part is to control the URLLC, which pursues super reliable and low latency requirements. Precisely, D2D fits within URLLC, especially in mission-critical communications. This article discussed the critical concept and advantages of NOMA techniques, which constitute promising future 5G systems. The principal NOMA schemes have been presented to compare their operating principles, key features, and receiver complexity. We also highlighted a range of critical challenges, opportunities and future research trends related to NOMA's design with eMBB, mMTC and URLLC. It is expected that NOMA will play an essential role in future 5G wireless communication systems sustaining massive connectivity and low latency.

In this paper, we investigated opportunistic spectrum sharing for a D2D-based URLLC system. A novel performance examined a NOMA-based D2D communication system coexisting with a cellular network, where the power resource requires to be carefully allocated based on the shortage in practice. We formulated a max–min fairness optimization problem with energy budget constraints to provide a reasonable boundary rate for the downlink to all devices and cellular users in the network while still limiting the whole system's total transmit power. Moreover, we consider two scenarios where a D2D system coexists with a mobile cellular system within the same cell. Two frequency bands are available, D2D users only use the first one, and D2D and cellular network users share the second one. D2D users use GADIA to allocate them to one of the two frequencies. GADIA is a fully distributed algorithm for finding a sub-optimal frequency band allocation to the network users. Each D2D group includes several paired devices so that the NOMA technique can be used to mitigate co-subchannel interference. Then, simulations in theoretical and practical scenarios are examined, comparing the maximum achievable rate and the EE at a given SE when using NOMA and OFDMA. The simulations were shown in MATLAB, and the results confirmed that NOMA obtains better performance than OFDMA in the categories compared. Furthermore, we study the mean interference received by D2D users by applying three different approaches, i.e., GADIA controlled frequency allocation, random frequency allocation, and without using any frequency sharing.

However, there are still diverse, challenging problems to be solved. Further challenges should also be determined in the context of NOMA systems, including the allied reference signal design and channel estimation, the loss of the PAPR in multi-carrier NOMA systems, such as LDS-OFDM, sustaining system scalability, the effects of channel-quality feedback design, BS cooperation, etc. Moreover, the existing multiple access design routinely allows

a single scheme for all applications, regardless of their distinct requirements. Therefore, different system design factors have to be estimated to accommodate the worst-case situation, which leads to inefficient multiple access design in various applications. Therefore, the software-defined multiple access technology is assumed to maintain the flexible configuration of multiple access schemes, and thus multiple services, as well as applications, can be advanced in 5G. It is anticipated that NOMA solutions will achieve further performance enhancements by addressing these challenges.

Abbreviations

The following abbreviations are used in this manuscript:

BS	Base station.
BW	Bandwidth of system.
UEs	Users/Terminals.
$v \in (1, 2, \dots, V)$	The number of sub-channel.
$u \in (1, 2, \dots, U)$	The number of user.
P_{BS}	The power at base station.
B_v	Bandwidth of the sub-channle v.
$P_{u,v}$	The power of the user at the sub-channel v.
x_v	The signal sent by BS.
s_i	The modulated symbol of the ith user.
$y_{j,v}$	The received signal of the jth user at the sub-channel v.
$h_{j,v}$	The coefficient of the sub-channel v from BS to the jth user.
$w_{j,v}$	Additive white Gaussian noise with mean of 0 and variance of σ^2 .
P_T	The base station transmit power.
α_k	is the power allocation factor for the UE_K power allocation factor.
n_{uk}	represents the received Gaussian white noise.
h_{bu}	indicates the channel from the base station to UE_K.
W	is the transmission bandwidth.
J	is the fairness index.
J'	is the target fairness index in the network.
EE	energy efficiency.
P_T	is the power consumed by the signal transmission.
P_s	is the power consumed due to other hardware and signal processing operations.

References

1. Series, M. Minimum Requirements Related to Technical Performance for IMT-2020 Radio Interface(s). Report ITU-R M.2410-0. 2017. Available online: https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2410-2017-PDF-E.pdf (accessed on 2 December 2021).
2. Mukherjee, A. Energy efficiency and delay in 5G ultra-reliable low-latency communications system architectures. *IEEE Netw.* **2018**, *32*, 55–61. [CrossRef]
3. Pedersen, K.; Pocovi, G.; Steiner, J.; Maeder, A. Agile 5G scheduler for improved E2E performance and flexibility for different network implementations. *IEEE Commun. Mag.* **2018**, *56*, 210–217. [CrossRef]
4. Popovski, P.; Nielsen, J.J.; Stefanovic, C.; De Carvalho, E.; Strom, E.; Trillingsgaard, K.F.; Bana, A.S.; Kim, D.M.; Kotaba, R.; Park, J.; et al. Wireless access for ultra-reliable low-latency communication: Principles and building blocks. *IEEE Netw.* **2018**, *32*, 16–23. [CrossRef]
5. Anand, A.; De Veciana, G.; Shakkottai, S. Joint scheduling of URLLC and eMBB traffic in 5G wireless networks. *IEEE/ACM Trans. Netw.* **2020**, *28*, 477–490. [CrossRef]
6. Chen, Z.; Sohrabi, F.; Yu, W. Sparse activity detection for massive connectivity. *IEEE Trans. Signal Process.* **2018**, *66*, 1890–1904. [CrossRef]
7. Ferrante, G.C.; Ostman, J.; Durisi, G.; Kittichokechai, K. Pilot-assisted short-packet transmission over multiantenna fading channels: A 5G case study. In Proceedings of the 2018 52nd Annual Conference on Information Sciences and Systems (CISS), Princeton, NJ, USA, 21–23 March 2018; pp. 1–6.
8. Kar, U.N.; Sanyal, D.K. An overview of device-to-device communication in cellular networks. *ICT Express* **2018**, *4*, 203–208. [CrossRef]
9. Elouafadi, R.; Benjillali, M. Cooperative NOMA-based D2D communications: A survey in the 5G/IoT context. In Proceedings of the 2018 19th IEEE Mediterranean Electrotechnical Conference (MELECON), Marrakech, Morocco, 2–7 May 2018; pp. 132–137.
10. Ngo, H.Q.; Ashikhmin, A.; Yang, H.; Larsson, E.G.; Marzetta, T.L. Cell-free massive MIMO versus small cells. *IEEE Trans. Wirel. Commun.* **2017**, *16*, 1834–1850. [CrossRef]
11. Van Chien, T.; Björnson, E.; Larsson, E.G. Joint power allocation and load balancing optimization for energy-efficient cell-free massive MIMO networks. *IEEE Trans. Wirel. Commun.* **2020**, *19*, 6798–6812. [CrossRef]
12. Yoon, T.; Nguyen, T.H.; Nguyen, X.T.; Yoo, D.; Jang, B. Resource allocation for NOMA-based D2D systems coexisting with cellular networks. *IEEE Access* **2018**, *6*, 66293–66304. [CrossRef]
13. Shafi, M.; Zhang, J.; Tataria, H.; Molisch, A.F.; Sun, S.; Rappaport, T.S.; Tufvesson, F.; Wu, S.; Kitao, K. Microwave vs. millimeter-wave propagation channels: Key differences and impact on 5G cellular systems. *IEEE Commun. Mag.* **2018**, *56*, 14–20. [CrossRef]
14. Wu, Q.; Chen, W.; Ng, D.W.K.; Schober, R. Spectral and energy-efficient wireless powered IoT networks: NOMA or TDMA? *IEEE Trans. Veh. Technol.* **2018**, *67*, 6663–6667. [CrossRef]
15. Chu, Z.; Yu, W.; Xiao, P.; Zhou, F.; Al-Dhahir, N.; ul Quddus, A.; Tafazolli, R. Opportunistic spectrum sharing for D2D-based URLLC. *IEEE Trans. Veh. Technol.* **2019**, *68*, 8995–9006. [CrossRef]
16. Zhang, H.; Liu, N.; Chu, X.; Long, K.; Aghvami, A.H.; Leung, V.C. Network slicing based 5G and future mobile networks: Mobility, resource management, and challenges. *IEEE Commun. Mag.* **2017**, *55*, 138–145. [CrossRef]
17. Ericsson, A. Common public radio interface (CPRI); interface specification v7.0. In *NEC Corporation, Alcatel Lucent, and Nokia Networks*; Huawei Technologies Co. Ltd.: Shenzhen, China, 2015.
18. Pfeiffer, T. Next generation mobile fronthaul and midhaul architectures. *J. Opt. Commun. Netw.* **2015**, *7*, B38–B45. [CrossRef]
19. Khorsandi, B.M.; Colle, D.; Tavernier, W.; Raffaelli, C. Adaptive function chaining for efficient design of 5G xhaul. In *International IFIP Conference on Optical Network Design and Modeling*; Springer: Cham, Switzerland, 2019; pp. 94–107.
20. Larsen, L.M.; Checko, A.; Christiansen, H.L. A survey of the functional splits proposed for 5G mobile crosshaul networks. *IEEE Commun. Surv. Tutor.* **2018**, *21*, 146–172. [CrossRef]
21. Guan, W.; Wen, X.; Wang, L.; Lu, Z.; Shen, Y. A service-oriented deployment policy of end-to-end network slicing based on complex network theory. *IEEE Access* **2018**, *6*, 19691–19701. [CrossRef]
22. Wu, Z.; Zhao, F.; Liu, X. Signal space diversity aided dynamic multiplexing for eMBB and URLLC traffics. In Proceedings of the 2017 3rd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, 13–16 December 2017; pp. 1396–1400.
23. Ranaweera, C.; Wong, E.; Nirmalathas, A.; Jayasundara, C.; Lim, C. 5G C-RAN with optical fronthaul: An analysis from a deployment perspective. *J. Light. Technol.* **2017**, *36*, 2059–2068. [CrossRef]
24. Lien, S.Y.; Hung, S.C.; Deng, D.J.; Wang, Y.J. Efficient ultra-reliable and low latency communications and massive machine-type communications in 5G new radio. In Proceedings of the GLOBECOM 2017—2017 IEEE Global Communications Conference, Singapore, 4–8 December 2017; pp. 1–7.
25. Chitimalla, D.; Kondepu, K.; Valcarenghi, L.; Tornatore, M.; Mukherjee, B. 5G fronthaul–latency and jitter studies of CPRI over Ethernet. *J. Opt. Commun. Netw.* **2017**, *9*, 172–182. [CrossRef]
26. Mahmood, N.H.; Lauridsen, M.; Berardinelli, G.; Catania, D.; Mogensen, P. Radio resource management techniques for eMBB and mMTC services in 5G dense small cell scenarios. In Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall), Montreal, QC, Canada, 18–21 September 2016; pp. 1–5.
27. Tian, L.; Yan, C.; Li, W.; Yuan, Z.; Cao, W.; Yuan, Y. On uplink non-orthogonal multiple access for 5G: Opportunities and challenges. *China Commun.* **2017**, *14*, 142–152. [CrossRef]

28. Giannone, F.; Gupta, H.; Kondepu, K.; Manicone, D.; Franklin, A.; Castoldi, P.; Valcarenghi, L. Impact of RAN virtualization on fronthaul latency budget: An experimental evaluation. In Proceedings of the 2017 IEEE Globecom Workshops (GC Wkshps), Singapore, 4–8 December 2017; pp. 1–5.

29. Gupta, H.; Manicone, D.; Giannone, F.; Kondepu, K.; Franklin, A.; Castoldi, P.; Valcarenghi, L. How much is fronthaul latency budget impacted by RAN virtualisation? In Proceedings of the 2017 IEEE Conference on Network Function Virtualization and Software Defined Networks (NFV-SDN), Berlin, Germany, 6–8 November 2017; pp. 315–320.

30. Nikaein, N. Processing radio access network functions in the cloud: Critical issues and modeling. In Proceedings of the 6th International Workshop on Mobile Cloud Computing and Services, MCS@MobiCom 2015, Paris, France, 11 September 2015; pp. 36–43.

31. da Rosa Righi, R.; Andrioli, L.; Rodrigues, V.F.; da Costa, C.A.; Alberti, A.M.; Singh, D. Elastic-RAN: An adaptable multi-level elasticity model for Cloud Radio Access Networks. *Comput. Commun.* **2019**, *142*, 34–47. [CrossRef]

32. Hasan, M.; Hossain, E.; Niyato, D. Random access for machine-to-machine communication in LTE-advanced networks: Issues and approaches. *IEEE Commun. Mag.* **2013**, *51*, 86–93. [CrossRef]

33. Laya, A.; Alonso, L.; Alonso-Zarate, J. Is the random access channel of LTE and LTE-A suitable for M2M communications? A survey of alternatives. *IEEE Commun. Surv. Tutorials* **2013**, *16*, 4–16. [CrossRef]

34. Ali, M.S.; Hossain, E.; Kim, D.I. LTE/LTE-A random access for massive machine-type communications in smart cities. *IEEE Commun. Mag.* **2017**, *55*, 76–83. [CrossRef]

35. Yang, W.; Wang, M.; Zhang, J.; Zou, J.; Hua, M.; Xia, T.; You, X. Narrowband wireless access for low-power massive internet of things: A bandwidth perspective. *IEEE Wirel. Commun.* **2017**, *24*, 138–145. [CrossRef]

36. Mahmood, N.H.; Abreu, R.; Böhnke, R.; Schubert, M.; Berardinelli, G.; Jacobsen, T.H. Uplink grant-free access solutions for URLLC services in 5G new radio. In Proceedings of the 2019 16th International Symposium on Wireless Communication Systems (ISWCS), Oulu, Finland, 27–30 August 2019; pp. 607–612.

37. Pocovi, G.; Soret, B.; Pedersen, K.I.; Mogensen, P. MAC layer enhancements for ultra-reliable low-latency communications in cellular networks. In Proceedings of the 2017 IEEE International Conference on Communications Workshops (ICC Workshops), Paris, France, 21–25 May 2017; pp. 1005–1010.

38. Shariatmadari, H.; Li, Z.; Uusitalo, M.A.; Iraji, S.; Jäntti, R. Link adaptation design for ultra-reliable communications. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–5.

39. Ohmann, D.; Awada, A.; Viering, I.; Simsek, M.; Fettweis, G.P. Diversity trade-offs and joint coding schemes for highly reliable wireless transmissions. In Proceedings of the 2016 IEEE 84th Vehicular Technology Conference (VTC-Fall), Montreal, QC, Canada, 18–21 September 2016; pp. 1–6.

40. Pedersen, K.I.; Khosravirad, S.R.; Berardinelli, G.; Frederiksen, F. Rethink hybrid automatic repeat request design for 5G: Five configurable enhancements. *IEEE Wirel. Commun.* **2017**, *24*, 154–160. [CrossRef]

41. Berardinelli, G.; Khosravirad, S.R.; Pedersen, K.I.; Frederiksen, F.; Mogensen, P. On the benefits of early HARQ feedback with non-ideal prediction in 5G networks. In Proceedings of the 2016 International Symposium on Wireless Communication Systems (ISWCS), Poznan, Poland, 20–23 September 2016; pp. 11–15.

42. Liao, Q.; Baracca, P.; Lopez-Perez, D.; Giordano, L.G. Resource scheduling for mixed traffic types with scalable TTI in dynamic TDD systems. In Proceedings of the 2016 IEEE Globecom Workshops (GC Wkshps), Washington, DC, USA, 4–8 December 2016; pp. 1–7.

43. Pedersen, K.I.; Niparko, M.; Steiner, J.; Oszmianski, J.; Mudolo, L.; Khosravirad, S.R. System level analysis of dynamic user-centric scheduling for a flexible 5G design. In Proceedings of the 2016 IEEE global communications conference (GLOBECOM), Washington, DC, USA, 4–8 December 2016; pp. 1–6.

44. Slalmi, A.; Saadane, R.; Chehri, A.; Kharraz, H. How will 5G transform industrial IoT: Latency and reliability analysis. In *Human Centred Intelligent Systems*; Springer: Singapore, 2021; pp. 335–345.

45. Zhu, L.; Xiao, Z.; Xia, X.G.; Wu, D.O. Millimeter-wave communications with non-orthogonal multiple access for B5G/6G. *IEEE Access* **2019**, *7*, 116123–116132. [CrossRef]

46. Park, S.; Truong, A.Q.; Nguyen, T.H. Power control for sum spectral efficiency optimization in MIMO-NOMA systems with linear beamforming. *IEEE Access* **2019**, *7*, 10593–10605. [CrossRef]

47. Xiong, C.; Li, G.Y.; Zhang, S.; Chen, Y.; Xu, S. Energy-efficient resource allocation in OFDMA networks. *IEEE Trans. Commun.* **2012**, *60*, 3767–3778. [CrossRef]

48. Ali, M.S.; Tabassum, H.; Hossain, E. Dynamic user clustering and power allocation for uplink and downlink non-orthogonal multiple access (NOMA) systems. *IEEE Access* **2016**, *4*, 6325–6343. [CrossRef]

49. Islam, S.R.; Zeng, M.; Dobre, O.A.; Kwak, K.S. Resource allocation for downlink NOMA systems: Key techniques and open issues. *IEEE Wirel. Commun.* **2018**, *25*, 40–47. [CrossRef]

50. Gau, R.H.; Chiu, H.T.; Liao, C.H. A geometric approach for optimal power control and relay selection in NOMA wireless relay networks. *IEEE Trans. Commun.* **2020**, *68*, 2032–2047. [CrossRef]

51. Fang, F.; Zhang, H.; Cheng, J.; Leung, V.C. Energy-efficient resource allocation for downlink non-orthogonal multiple access network. *IEEE Trans. Commun.* **2016**, *64*, 3722–3732. [CrossRef]

52. Dai, L.; Wang, B.; Yuan, Y.; Han, S.; Chih-Lin, I.; Wang, Z. Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends. *IEEE Commun. Mag.* **2015**, *53*, 74–81. [CrossRef]

53. Babadi, B.; Tarokh, V. GADIA: A greedy asynchronous distributed interference avoidance algorithm. *IEEE Trans. Inf. Theory* **2010**, *56*, 6228–6252. [CrossRef]

54. Al-Abbasi, Z.Q.; So, D.K.C. User-Pairing Based Non-Orthogonal Multiple Access (NOMA) System. In Proceedings of the 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), Nanjing, China, 15–18 May 2016; pp. 1–5. [CrossRef]

55. Guo, J.; Wang, X.; Yang, J.; Zheng, J.; Zhao, B. User Pairing and Power Allocation for Downlink Non-Orthogonal Multiple Access. In Proceedings of the 2016 IEEE Globecom Workshops (GC Wkshps), Valencia, Spain, 6–8 June 2018; pp. 1–6. [CrossRef]

56. Zhang, H.; Zhang, D.; Meng, W.; Li, C. User pairing algorithm with SIC in non-orthogonal multiple access system. In Proceedings of the 2016 IEEE International Conference on Communications (ICC), Kuala Lumpur, Malaysia, 22–27 May 2016; pp. 1–6. [CrossRef]